DOCUMENT RESUME

ED 162 863 SE 025 358

AUTHOR	Osborne, Alan R., Ed.
TITLE	Investigations in Mathematics Education, Vol. 11, No. 2, Spring 1978.
INSTITUTION	Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE	78
NOTE	73p.; Contains occasional light and broken type
AVAILABLE FROM	Information Reference Center (ERIC/IRC), The Ohio State Univ., 1200 Chambers Rd., 3rd Floor, Columbus, OH 43212 ($6.00 subscription, $1.75 ea.)
EDRS PRICE	MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS	*Abstracts; *Cognitive Ability; *Effective Teaching; Elementary Secondary Education; Higher Education; Instruction; Learning; *Mathematics Education; Problem Solving; *Research Reviews (Publications); Student Characteristics; Teacher Characteristics; Teacher Education; *Teaching Methods; Testing

ABSTRACT

Twenty research reports related to mathematics education are abstracted and analyzed. Five of the reports deal with various student characteristics, three with cognitive ability, two with teacher characteristics, two with effective teaching, seven with teaching methods and one each with testing and problem solving. Research related to mathematics education which was reported in RIE and CIJE between January and March 1978 is listed (MP)

ED162863

INVESTIGATIONS IN MATHEMATICS EDUCATION

Published quarterly by

The Center for Science and Mathematics Education
The Ohio State University
1945 North High Street
Columbus, Ohio 43210

With the cooperation of the ERIC Clearinghouse for Science, Mathematics, and Environmental Education

Volume 11, Number 2 - Spring 1978

Subscription Price: $6.00 per year. Single Copy Price: $1.75
Add 25¢ for Canadian mailings and 50¢ for foreign mailings.

SE 025 358

Spring 1978

RELATIONSHIPS BETWEEN SELECTED TEACHER BEHAVIORS OF PREALGEBRA TEACHERS AND SELECTED CHARACTERISTICS OF THEIR STUDENTS. Campbell, N. Jo; Schoen, Harold L. Journal for Research in Mathematics Education, v8 n5, pp369-375, November 1977.

Expanded Abstract and Analysis Especially Prepared for I.M.E. by James M. Moser, University of Wisconsin-Madison.

1. Purpose

The purpose of the study was to determine whether relationships exist between students' perceptions of selected behaviors of prealgebra teachers and selected student performance and attitude variables.

2. Rationale

Little rationale was presented, other than brief reference given to use of students' opinions by three other researchers (Murray, 1972; Ryans, 1953; and Cogan, 1958).

3. Research Design and Procedure

Data were gathered in Fall 1974 by means of a questionnaire that was validated by three faculty members and subsequently revised after a pilot test with three prealgebra classes (at the junior high level before ninth-grade algebra). Items included those collecting information about students' characteristics (grades in mathematics and in other subjects) and students' attitudes towards mathematics, their mathematics teacher, and school in general, as well as the 19 items dealing with students' perceptions as to whether teachers exhibited a variety of teacher behaviors. These latter items were rated on a four-point scale: 0--can't answer; 1--never; 2--sometimes; and 3--always.

The subjects were 816 female and 786 male students drawn from prealgebra classes of 73 teachers in 28 different school districts in six southeastern Oklahoma counties. Each teacher selected one class for inclusion in the study. The mean grade level for the 73 classes was 7.90. The responses from several randomly selected pairs of students from different classes were correlated to obtain estimates of interrater reliability. These estimates ranged from .65 to .83.

Class mean responses were computed for each item with corrections made for negatively worded statements. Intercorrelations for class means of students' characteristics and students' attitudes were computed and tested for significance at the .005 level. Correlations between characteristics and attitudes of students and the 19 teacher-behavior items were also computed and tested for significance at the .0005 level. Student responses to items about teaching behaviors were factor-analyzed using the truncated component model. After six factors were identified, class means for these factors were computed and then correlation coefficients computed between these mean scores and the mean scores of the characteristics and attitudes of the students.

4. Findings

Attitudes of students toward mathematics were positively related with their grades in mathematics, with the comparison of their mathematics grades to other grades, and most strongly with their attitudes toward their mathematics teacher. Their attitudes toward their mathematics teacher were positively related to their comparisons of grades in mathematics to their other grades. Of all the correlations computed between characteristics and attitudes of students and the teacher-behavior items, only one was determined to be statistically significant--between attitude towards the mathematics teacher and the behavior of "shows continuity of the mathematics curriculum."

The six identifiable factors of teacher behavior that had eigen values greater than 1 after using the Varimax rotation procedure were labeled as (1) positive teaching orientation, (2) flexible teaching methods, (3) lack of concern for student growth, (4) traditional teaching orientation, (5) use of physical models, and (6) orientation toward students. Of the 30 correlation coefficients computed between these six factors and the five student characteristic and attitude scores, only one was found to be significant--between attitude toward teacher and the first factor.

5. Interpretations

The study identified several moderate relationships between student perceptions of selected teacher behaviors and student achievement or attitudes. The teacher behaviors that most often tended to be related with positive student attitudes toward mathematics and the mathematics teacher were, "Explains <u>why</u> with <u>how</u> problems are worked" and "Shows continuity of the mathematics curriculum." Students who perceived their mathematics teacher as trying to remove the "mysteries" of mathematics had more positive attitudes toward mathematics and the teacher.

Critical Commentary

One of the first questions that comes to mind is the one that comes with any study of teaching behavior at any level: namely, "How were the teachers selected?" There is no mention of the selection process in the research report. It is lamentable that the classes of the 73 teachers were not randomly chosen by the researcher. One fears that the results are biased as a result. The reported mean grade level of 7.90 certainly indicates that there is a bias toward older children, since it is assumed that (since no specific mention is made of the fact) the majority of classes were seventh and eighth graders.

Another major problem is the scaling of the student responses on the characteristic and attitude items. The responses appear to be forced into a three-point scale that lacks the desirable psychometric properties, given that so many correlations were computed. For example, on the attitudes toward mathematics, mathematics teacher, and school in general, the choices were "like very much" {italics mine}, "OK", and "don't like." The gap between the neutral and positive responses seems much greater than between the netural and negative responses. This is especially crucial in the attitude toward teacher item; there are probably some severe reliability problems here! The mathematics grades choices seemed a bit strange as well--"A's and B's", "B's and C's," and "C's and D's." Would a child who had 1 A and 4 B's receive the same score as the child who had 4 A's and 1 B? I hope not, but on the basis of the report I think it is the case.

It is interesting to note the limited rationale presented. Only three brief references were given, two of which are over 15 years old. This does not seem to be a very active field of investigation! The lack of randomization, the presence of data based upon questionable measurement scales, and the presence of very few "moderate" relationships suggests that this study has very limited applicability and is of interest to a limited number of professionals.

THE RELATIVE EFFECTIVENESS OF THREE GEOMETRIC PROOF CONSTRUCTION STRATEGIES. Carroll, Dennis C. Journal for Research in Mathematics Education, v8 n1, pp62-67, January 1977.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Arthur F. Coxford, University of Michigan.

1. Purpose

The purpose was to determine the relative effectiveness of three strategies of geometric proof construction: analysis, synthesis, and combined analysis-synthesis.

2. Rationale

The three strategies identified above are frequently referred to in methods books and employed in different school geometry textbooks. Available evidence has not identified an optimal strategy for use with school geometry students. No theoretical, psychological, or other basis was indicated.

3. Research Design and Procedure

Nine intact classes were chosen from five high schools to study a six-day, experimenter-developed unit on congruent triangles. The instruction was performed by nine student teachers especially trained in the three treatments. The nine student teachers were randomly assigned to treatment, three for each treatment. The experimental unit was taught approximately 12 weeks into the first semester. Previously, the students had studied geometry using Modern Geometry (Houghton Mifflin, 1975) as a text.

Upon completion of the six days of instruction, two tests of proof construction (RGT and EGT) achievement were administered. The RGT (required-given test) was made up of five items randomly selected from a pool of ten items. Each item contained a figure, the given relating to the figure, and the deducible statment. In the given, only necessary information was included. The EGT (extraneous given test) was similar to the RGT with the exception that the given included extraneous information. During instruction, approximately one-half of the proof practice problems contained extraneous information in the given; thus, the EGT was not a transfer test. The reliability coefficients of RGT and EGT were 0.92 and 0.93, respectively.

The completed tests for each class were evaluated by the student teacher, by the regular classroom teacher, and by the experimenter. The scoring scale was 0 to 3 for each item, with a maximum of 15 points for each test. The three evaluators' scores were averaged to obtain each student's score. A student was categorized as above average or below average on the basis of the student's first six weeks grade in geometry and the grades for first- and second-semester algebra.

4. Findings

For each strategy group, the mean scores on the RGT and the EGT for the above- and below-average groups were data points. These data were submitted to a 3x2 multivariate analysis of variance for identification of a composite of the dependent variables and effect testing. The composite was the difference in the achievement levels on the RGT and the EGT. For the three strategy groups the composite was: analysis -- 2.23; synthesis -- 0.51; combination -- 0.41. It was found that the mean composite scores of the strategy groups were different ($p < .05$) and that the mean composite scores of the above- and below-average prior-achievement groups were different ($p < .005$).

Univariate analysis of variance failed to show that the strategy groups' mean RGT or EGT scores were different. Post hoc analysis showed that the mean composite score of the analytic strategy group was greater than either strategy group composite score. No other differences were found.

5. Interpretations

The major finding was that the analytic strategy group showed a large decrease in proof construction achievement when encountering extraneous data in the given, while the other groups decreased only slightly. The author suggested that this result might imply a modification in the Bechtold and Scandura contention that a reduction in achievement occurs when extraneous data are included in a problem-solving situation. Such a general statement may need qualification and refinement in terms of other pertinent variables.

Since the analytic strategy was less stable across problem types, the author suggested that the synthetic or combination strategy be employed in geometry courses until definitive evidence is found for an optimal strategy. The author noted the limitations of the small sample, the experience of the teachers, and the brief instructional period.

Critical Commentary

The author should be applauded for attempting to develop knowledge in an area as complex as proof strategies. Even though there are many limitations on this study, it is a tentative step in helping teachers to do a more effective job in teaching proof. I would like to see additional studies that focus on the variables affecting proof construction achievement rather than studies seeking an "optimal" strategy. The former focus would be far more manageable and free from extraneous forces.

THE ELEMENTARY SCHOOL AS A TRAINING LABORATORY AND ITS EFFECT ON LOW-ACHIEVING SIXTH GRADERS. Fennell, Francis and Trueblood, Cecil. Journal for Research in Mathematics Education, v8 n2, pp97-106, March 1977.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by James M. Sherrill, University of British Columbia.

1. Purpose

"...to collect formative data to show the impact of two teacher-training experiences on low-achieving sixth-grade pupils."

2. Rationale

The traditional means-referenced conception of instruction has been questioned as being inadequate for most instructional decision making. In its place several leading teacher educators have suggested a goal-referenced instructional model. The adoption of a goal-referenced instructional model poses new problems; specifically, how should teacher education programs be designed to prepare prospective elementary teachers to perform the many professional tasks demanded by the emerging types of individualized instruction?

3. Research Design and Procedure

One group of five elementary teacher education students (TES1) was trained in the tasks associated with a goal-referenced model of instruction; a second group (TES2) was trained to implement the procedures set forth in the teacher's guide for the classroom text assigned by a local school district. Both TES1 and TES2 were given the same number of class hours of instruction and review of the mathematics content to be taught. A set of behavioral teaching competencies was used with both groups to evaluate the method instructor's performance and to insure, as much as possible, that both TES1 and TES2 had successfully completed their training programs.

Both groups of TES taught sixth-grade students judged as low achievers by their fifth-grade teacher and their scores on Form W (May 1971 administration) of the Stanford Achievement Test. Of the 47 pupils (N(TES1) = 20, N(TES2) = 27) in the study, 60 percent scored below the fifth stanine on all subtests. Only two of the subjects scored above the sixth stanine on any subtest. Due to school policy the pupils could not be randomly assigned to treatment groups.

The formative data include (a) pupils' pretest, posttest, and retention test scores on a 32-item unit mastery test; (b) pupils' pretest, posttest, and retention test scores on the Suydam-Trueblood Attitude Toward Mathematics scale; and (c) the amount of time needed by the instructional group to demonstrate mastery of the unit objectives. The unit mastery test was developed from a table of specifications based on the behavioral objectives for the unit of instruction. On a pilot administration of the unit mastery test, the test-retest reliability was 0.89. The Suydam-Trueblood attitude scale has an average internal consistency (Cronbach's Coefficient Alpha) of 0.96.

A single classification analysis of variance with repeated measures was computed for each instructional group to assess achievement gain. A Newman-Keuls analysis was used to locate significant differences. The single classification analysis of variance and Newman-Kuels analysis were used with the Suydam-Trueblood Attitude Toward Mathematics scale scores.

The unit of instruction for the study was functions and equations.

4. Findings

The ANOVA results showed that both groups of subjects made significant gains on the 32-item unit mastery test. The Newman-Keuls analysis showed that all three pairwise comparisons for pretest, posttest, and retention test were significant for TES1. Only the posttest-retention test failed to be significant comparison for TES2.

The ANOVA and Newman-Keuls analyses showed a significant gain in attitude for the TES1 group on the pretest-retention test comparison only. The ANOVA for TES2 pupils' attitude scores was not significant.

Every TES2 subject took nineteen 45-minute instructional periods to complete the unit. Half of the TES1 subjects finished the unit in 16 periods and one-fifth of the TES1 subjects finished the unit in 10 periods.

5. Interpretations

Both TES groups significantly increased their pupils' unit mastery test scores without negatively affecting their attitude toward elementary school mathematics. Therefore, using the elementary school as a training laboratory for the TES in this study seemed to have positive impact on the pupils they taught.

The data from this study suggest certain ideas that have potential application to the field-based preparation of prospective elementary teachers:

1. Having TES help teach an individually prescribed unit of mathematics under the supervision of a master teacher may be educationally defensible in terms of the positive impact it can have on the pupils and the realistic training experience it can afford the prospective teachers.

2. Public schools and teacher education institutions could consider combining forces to assist those pupils who most need individualized attention--the low-achieving pupils.

Hence, a more detailed and more carefully designed set of studies should be conducted.

Critical Commentary

The concerns about this study center around two areas: the original design of the study and the statistical analysis.

1. Design of the study: It is agreed that the study "...was not intended as an experimental comparison to decide which training procedure was better. However, a comparison group was needed. While both TES groups increased the unit mastery test scores of their pupils, it may be that neither performed as well as a traditional non-field-based TES group might have performed when they started working with students. The gains may have been due to the confounding effects of maturation and testing. The study, as designed, suffers from a classical case of statistical regression effect.

2. Statistical analysis: Putting aside the facts that there was no random assignment and the authors admit that the independence-of-data issue is cloudy, there is still some confusion of the purpose of comparing pretest-posttest scores. One of the pieces of data collected was "...the amount of time needed by the instructional groups to demonstrate mastery of the unit objectives." It must be assumed that the researchers had some manner, other than the unit mastery test, to judge whether or not the individual students had attained mastery. The subjects were pretested, worked on the mathematics content until they were judged to have mastered the unit, then tested to see if they have mastered the unit. It could not have come as a surprise that the F ratios based on the unit mastery test data were very large.

In spite of the above criticisms the conclusions of the authors are realistic and stated with cautionary words such as "may" and "could". Using the same cautions, it may be possible that the study could serve as an existence statement for the authors' point of view.

SEX-RELATED DIFFERENCES IN MATHEMATICS ACHIEVEMENT AND RELATED FACTORS: A FURTHER STUDY. Fennema, Elizabeth H.; Sherman, Julia A. Journal for Research in Mathematics Education, v9 n3, pp189-203, May 1978.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Mary Grace Kantowski, University of Florida.

1. Purpose

To continue the study of cognitive and affective variables that influence males and females to learn mathematics at different levels. Cognitive variables include: computational skill, knowledge of concepts, problem solving ability, verbal ability, and spatial visualization. Affective variables include: attitude toward success in mathematics, mathematics as a male domain, perceived attitude of parents and teachers, effectance motivation, confidence, and usefulness.

2. Rationale

Several recent studies have suggested that the widely held belief of male superiority in mathematics is not as prevalant as had been believed and is, moreover, age-related. In earlier reported studies including only grades 9 through 12, the authors found that sex-related differences were found in only half of the school population sampled when the number of years of studying mathematics was controlled. This research is a follow-up study designed to look at the same variables in the feeder schools for those used in the studies reported earlier.

3. Research Design and Procedure

A battery of tests was administered to 1320 sixth-, seventh-, and eighth-grade students in middle schools in Madison, Wisconsin that were the feeder schools for the population of the previous study. The sample included only students in the top 85% in mathematics achievement. The tests included the Romberg-Wearne Problem Solving Test, vocabulary tests from the Verbal Battery of the Cognitive Abilities Test, the Space Relations Test, Mathematics Concepts Test, Mathematics Computation Test, and the Fennema-Sherman Mathematics Attitude Scales.

Means of the 15 measures were computed for males and females in each of the three grade levels and for each of four areas of the city. An ANOVA was performed on each variable, with sex, grade, and area used as sources of variance. Correlation coefficients between measures were computed for each sex and for the students combined over area and grade. A principal component factor analysis was also performed on all variables combined over area and grade.

4. Findings

Means and standard deviations for all measures were reported by area, sex and grade, in addition to the usual F ratios of ANOVAs (Sex x Grade x Area, and Sex x Grade for each Area). Significant sex-related differences were found in only two affective variables, in each case "favoring" the male: "Confidence in Learning Mathematics" and "Mathematics as a Male Domain." As expected, significant area effects were found for all variables, and significant grade effects were found for the cognitive variables.

The results of the Sex x Grade for each Area data analysis showed the following: (1) significant differences in all areas for "Mathematics as a Male Domain"; (2) significant difference in Computation (favoring females) in Area 4; (3) significant differences in favor of males in Romberg-Wearne Application and Romberg-Wearne Problem Solving and for six of the eight affective variables in Area 3. Only the "teacher" and "Effectance in Motivation" variables showed no significant sex differences in Area 3.

5. Interpretations

The findings "strongly suggest that there are no universal sex-related differences in mathematics learning." The authors note that the results of this study agree with the NAEP results of lack of differences in mathematics achievement before age 17, but are in conflict with the NLSMA conclusions that males are superior on tasks of high cognitive complexity. They suggest that the heightened interest in women in mathematics in the interim years could be at least partially responsible for some of the differences in results found in NLSMA studies and in this one.

One surprising result was the lack of significant difference in spatial visualization in males and females, a finding that would, if further substantiated, dispel the long-held belief that males are superior to females in spatial ability.

A very interesting aspect of the discussion of results is the comparison of the findings of this study with those of the above-mentioned study in grades 9 through 12. This is particularly true where sharp differences were observed in the affective measures. Especially noteworthy are the "Confidence in Learning Mathematics" and the "teacher" variables and the relationship between these variables.

Critical Commentary

The Fennema-Sherman studies are a valuable contribution to the search for reasons for the dearth of women in mathematics-related fields. More such well-designed and carefully conducted research is needed to provide hard data to substantiate hypotheses or to dispel popular myths.

In addition to presenting status information, these studies provide a baseline with which to compare the results of future studies.

A careful study of the tables and discussion suggests the following questions and comments:

(1) How much a function of Madison, Wisconsin are the results? Although a socioeconomic mix does exist, any university town is an atypical sample. Comparative studies are needed.

(2) What are the socioeconomic characteristics of Areas 1, 2, and 4? It would help the reader to have some demographic information on each of the areas since one purpose of such studies is to provide information to support hypotheses for probable causes of lack of participation and for achievement differences.

(3) The lack of significant sex-differences in the spatial test was an especially interesting finding in need of further investigation. Since it is generally accepted that more than one space factor exists, further substantiation of the findings with other measures of spatial ability are indicated.

(4) Some of the graphs on page 199 of the article are misleading. A perusal of the table of means (pp. 192-193) suggests some interesting discrepencies. The "Confidence in Mathematics " graph would lead the leader to believe that female confidence was consistently lower. In fact this was not the case in four of the nine classes studied. Likewise, the "Usefulness of Mathematics" graph shows males consistently higher. Yet the means for females in five of the nine classes are higher. As the author noted (p. 198), great differences in favor of males especially on the affective variables occurred in Area 3. These large discrepancies in one area could account for an inaccurate picture.

(5) The high correlations between the students' confidence in mathematics and their perceived attitudes of parents and teachers toward them as learners of mathematics should provide hypotheses for further study.

FLOW CHARTS IN MATHEMATICS CLASSES FOR ELEMENTARY SCHOOL TEACHERS. Ford, Janet E. and McLeod, Douglas B. Two-Year College Mathematics Journal, v8 n1, pp15-19, January 1977.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Otto C. Bassler, George Peabody College for Teachers.

1. Purpose

To develop a unit on flow charts and to determine its effectiveness in helping teachers explain some algorithms from arithmetic.

2. Rationale

Recent recommendations by several committees and mathematics educators have indicated that flow charts can be a useful device in teaching mathematical concepts. One particular topic that is suited to the use of flow charts is the development of algorithms in arithmetic. It is conjectured that a unit on flow charts would be helpful to students as they identify steps in an algorithm as well as helping them justify algorithms.

3. Research Design and Procedure

A unit of instruction was designed to teach flow charts and to explain arithmetic algorithms using flow charts. The development of the unit was accomplished by specifying the objectives, writing and sequencing instructional activities to attain these objectives, and revising and improving the unit based upon the results of three pilot studies.

Twenty-four female students, enrolled in a first semester mathematics course for elementary teachers, were assigned at random to two treatment groups. The experimental group learned flow charts and then used flow charts to study arithmetic algorithms. The control group studied the same algorithms without using flow charts. Both groups used the same manipulative materials in instruction that emphasized identifying the steps in algorithms and justifying algorithms. Each of the two investigators was randomly assigned three hours to teach each treatment group, resulting in a total of six instructional periods for each group. With the exception of the flow charts that were only taught to the experimental group, the same algorithms, problems and exercises were used in both groups.

Following the treatments both groups were given a post-test on three algorithms--two of which had been discussed during instruction and one which had not been discussed in either treatment group. The form of the test required students to complete five examples of the algorithm, then to write a list of instructions for the algorithm, and also to explain why their list of instructions produced the correct answer. Two dependent variables, "list steps in algorithm" and "justify algorithm", were scored for each student's post-test. The maximum score for each dependent variable was 12, four points for each of the three algorithms. A brief attitude-toward-flow-charts questionnaire was also administered to the subjects in the experimental group.

4. Findings

Means of post-test scores were compared using t-tests. The results indicated that students who had studied flow charts performed significantly better ($p < .01$) on listing the steps involved in algorithms. Both groups performed at about the same level on justifying algorithms. The mean score for each group was quite low when justifying algorithms--1.8 out of 12 for the experimental group and 1.9 for the control group. The results of the attitude survey indicated generally positive attitudes of the students in the experimental group toward flow charts.

5. Interpretations

It was concluded that constructing flow charts did help students give a more complete listing of the steps in an algorithm. This is interpreted to offer some support for the inference that a unit on flow charts can help prospective teachers do a better job of explaining how algorithms work. The superiority of the experimental group was due to the ability of these subjects to describe all of the cases of an algorithm whereas subjects in the control group tended to use only one problem as a basis for their list of instructions. Neither group did well in justifying algorithms and it was concluded that flow charts do not seem to help students explain why algorithms give the correct answer.

Critical Commentary

The unit on flow charts appeared to have been well prepared following an appropriate curriculum development model. The authors are to be commended for revising the experimental treatment materials, based on the results of three pilot studies. No indication was provided about the instructional program for the control subjects other than it developed the same algorithms without the use of flow charts. Perhaps if the same curriculum development model had been applied to the control treatment, the results would have been different. There is no way for the reader to know the emphasis placed upon listing steps in the control treatment.

Both groups failed to achieve the objective, "justify algorithms", since students in both treatments had low mean achievement (about 15%) on this measure. Achievement this low would tend to suppress any differences that might exist between the treatments. If this was to be one of the outcomes of instruction, then the curriculum development model did not achieve its goal and the instructional materials need to be analyzed and revised.

No indication of the scoring scheme used to rate the dependent variables was provided nor were any estimates of test or rater reliability given. These conditions may have a substantial bearing on the outcomes of the study.

Finally it seems hazardous to generalize from a significant finding on listing the steps in an algorithm to doing a better job of explaining how algorithms work.

INDEFINITE GOALS IN WELL STRUCTURED PROBLEMS. Greeno, James G. Psychological Review, v83 n6, pp479-491, November 1976.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Max S. Bell, University of Chicago.

1. Summary

Excellence in "problem solving" is often said to be one of the central goals of school mathematics teaching, but we do not agree among ourselves about what that phrase means. "Problem solving" occupies many researchers in psychology but there seems little reason now to revise a psychologist's judgment of over a decade ago that "Research in human problem solving has a well-earned reputation for being the most chaotic of all identifiable categories of human learning" (Davis, 1966). Greeno's article uses school content as the vehicle for a problem-solving inquiry published in a well-known psychological journal. Hence, it should be a good place to test the actual overlap between our concern for teaching problem solving and the efforts of information processing theorists within psychology to illuminate how human beings go about problem solving.

Information processing inquiries typically aim to sort out fairly complex thinking processes by models that in concept or in actuality can be simulated with computer programs (Newell and Simon, 1972), and this article is in that tradition. The main thrust of the article is that open-ended or indefinite problem-solving goals or subgoals can arise in otherwise "well-structured" problems and not merely from the uncertainties of ill-structured problems. Definition of the phrase "well structured problem" with careful use of "problem," "problem state," "elements," "relations," "operators," "problem goal," and so on, takes up the first two pages of the article, along with the footnote warning that even so "there is no general definition of a well-structured problem, nor should there be" (p. 479). There can, however, be precise definition of some words, for example: "I will say that a goal is indefinite when its description in disjunctive normal form has at least two terms consisting of single features or conjunctions and when each such term has one or more features that are not present in other terms" (p. 480).

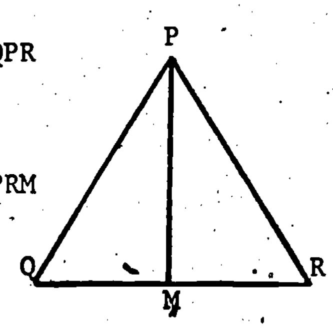
Next, by having five high school students talk through geometry problems such as the one exhibited here, Greeno shows that it is typical for students to go through an information gathering stage (e.g., by marking the figure) without a definite theorem in mind and only then begin to select specific strategies. In three pages of discussion this process is established as an example of a well-structured problem solved by generating indefinite subgoals.

$\overline{PM} \perp \overline{QR}$

$\overline{PM}$ bisects $\angle QPR$

Prove:

$\triangle PQM \cong \triangle PRM$



In the next four pages a computer program (Perdix) for doing simple congruence proofs in geometry is described in some detail as a model for indefinite subgoals used in solving well-structured problems. It is not possible to summarize that discussion briefly, but those familiar with such models in the information processing literature can get its flavor from the flow chart fragment shown in the article. The problem solving consists of a series of passes through such a process, trying to fit what is known at each pass to SAS, ASA, AAS, or hypotenuse-leg, in that order. If enough direct information is "given" in the problem as stated, then the proof will emerge during the first pass. If not, an "infer congruent parts" production routine is activated to get, if possible, just one additional congruent pair of sides or angles to work with, then the process is scanned again. If there is still not a solution, then "infer congruent parts" operates again, another pass is tried, and so on. The routine as programmed leads, if at all, to a single unique solution, even though a human problem solver might readily see several possibilities. It is clear that the richness and versitility of the "infer congruent parts" routine pretty much determine how complicated can be the problems handled by Perdix. Given the purposes of the present exposition, little detail is given about that routine or Perdix more generally. Lacking such detail, one cannot tell if a couple of apparent flaws in the partial flowchart given in the article are real, misprints, or attended to in ways not explained here.

The point, of course, is not to produce a computer program to do proofs so that humans will not have to do them but rather to try to model and thus illuminate what might be actual human thinking processes. The author explains some ways in which he thinks Perdix proceeds just as humans might and some reasons to believe that the program models use of indefinite goals in solving well-structured problems. There are two pages of interest primarily to specialists where the features of Perdix are compared to General Problem Solver (GPS) computer simulations of human thinking and to other computer simulations of congruence proofs. The article ends by making explicit the central point of the paper that "The ease with which the theory of well-structured problems can apparently be extended to accommodate indefinite goals seems encouraging for the possibility that other sources of uncertainty in ill-structured problems might also be incorporated in the theory of well-structured problems" (p. 491).

2. Commentary

As I turn to comment, it must be understood that I have no doubt about the worth and usefulness of the article for its primary audience of psychologists. Given that, it is still worth exploring whether it is equally useful to us. We seem often to assume that we are somehow remiss in not making direct use of the treasures of knowledge available from theoretical psychology models for the improvement of mathematics education, but our sense of guilt about that may not be warranted. Whatever the potential of such relatively pure research and model building for eventually increased understanding, there may still be serious and possibly fatal barriers to its direct application within our field.

The first of these difficulties is simply in communication between the two fields, mathematics education and psychology. On the one hand, only those among mathematics educators who immediately resonate to the acronym GPS and know about at least some of ACT, EPAM, CLS, or PLANNER will find this article easily accessible, since those acronyms and a variety of other special terms and references are used without further elaboration. On the other hand, the detailed explanations and protocol apparently needed to communicate to psychologists about how youngsters may go about simple congruence proofs may seem trivial to anyone who has taught a tenth-grade geometry course. That combination of obscurity and triviality cannot be a criticism of this article--which might well have been given the opposite tilt if written for us instead of psychologists--but it does serve to illustrate the communications problem.

Second, try as I might, I cannot see such a program as Perdix as truly modeling human problem solving in attacking such proofs as are used as examples. A sensible tenth grader would, as the protocol given in the article indicates, simply mark on the figure the parts given as or easily shown to be congruent, look at that to see what congruence theorem applies, and then undertake to fill in the details to make it a respectable exposition. He would not, as Perdix does, use only what is <u>directly</u> given to check out methodically SAS, then ASA, et cetera, and only then go looking for some <u>single</u> additional congruent pair in order to go through another series of triangle congruence tests. That is, I find it a little far-fetched to say that "the system as programmed may be quite realistic for the problem of proving congruence of triangles" (p. 488) where "realistic" apparently refers to how humans actually attack such a problem. The trouble as I see it is that such a program as Perdix ultimately comes down to an <u>algorithmic</u> process for obtaining single solutions for a restricted range of problems, while what we want to teach and what we want youngsters to use are heuristic approaches that encompass multiple solutions for a broad range of problems. That algorithmic versus heuristic distinction has been neatly expressed by L. N. Landa who has also demonstrated a specific heuristic routine that resulted in remarkable gains in Russian eighth graders' ability to prove geometry theorems (Landa, 1975). I can no more imagine the teaching of a Perdix-like algorithm as a fruitful means to similar gains than I can imagine a student hitting on a Perdix-like routine as a "natural" approach to doing geometric proofs. Again, this article does not suggest Perdix as a guide to instruction, so the remarks above are intended only to emphasize the considerable distance between such modeling of "problem solving" and the direct applicability to mathematics education. It should be said that the information processing theorists are quite aware that problem solving often cannot be algorithmic and that some of their computer simulations attempt to model use of heuristic routines.

A third and related barrier to application of such models is the gap between the precise and narrow limits that must be imposed to get clean results and the broader range of concerns typical in school instruction. For example, this article goes to some lengths to define a "well-structured problem" then works toward a certain extension of the theory of such problems. But I find it a little discouraging to learn that the theory has not already encompassed that extension, which most teachers would recognize as essential even in the relatively stylized problem of putting together a geometric proof.

To close on a positive note, one admires the inventiveness that enables geometry problems to be presented to computers in symbolic form and solutions achieved by general routines such as those of Perdix. More important, details of that may tell us something useful about how our somewhat sloppy heuristics of extracting information from diagrams could be enriched by exploiting our diagram labeling systems, say in confusedly overlapping figures. (It is not clear in the article whether Perdix can handle such "figures".) Some of that is already done in many books now, but perhaps more could be done. Again, it may be that details about the "infer congruent parts" routine would tell us something useful about how much information needs to be stored away to support richness in that routine.

The methodology of the inquiry itself can perhaps enrich our view of what is respectable in scholarly inquiry. Greeno has thought about a category of problems, observed some youngsters at work on them, formulated a conception, produced a routine to test ~~it~~ out, and then reported the results with barely any empirical data and no statistics to speak of. At a more "practical" level, our profession abounds in analogous opportunities to sort out specific instructional problems (not usually with computer routines), yet the doctoral students who produce the bulk of research literature seldom work in this way. Perhaps psychological research can at least teach more of us that there is a wider range of ways to do scholarly work on instruction and learning problems than is common in our field of mathematics education.

References

Davis, G. A. Current status of research and theory in human problem solving. Psychological Bulletin, v66, pp36-54, 1966.

Landa, L. N. Some problems of algorithmization and heuristics in instruction. Instructional Science, v4, pp99-112, 1975.

Newell, A. and Simon, H. A. Human problem solving. Englewood Cliffs, N.J.: Prentice-Hall, 1972.

A COMPARISON OF TWO APPROACHES TO THE ASSESSMENT OF CONDITIONAL-REASONING ABILITIES. Jansson, L. C. Journal for Research in Mathematics Education, v9 n3, pp175-188, May 1978.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Larry Sowder, Northern Illinois University.

1. Purpose

To compare "adolescents' abilities to handle simple conditional arguments as measured by two different assessment procedures."

2. Rationale

Researchers have used different instruments to assess conditional reasoning. Do different instruments measure the same thing?

3. Research Design and Procedure

Subjects were 185 eighth graders, 140 tenth graders, and 139 twelfth graders, all in mathematics classes.

First Instrument: In group settings, each student received, randomly, one of two 32-item investigator-constructed tests of conditional reasoning. These forms presented two premises and asked for a yes/no/maybe response to the truth of a given conclusion. Each form contained eight items based on each of these principles: (a) modus ponens (if p then q; p; therefore q); (b) if p then q; q; therefore p (invalid); (c) if p then q; not p; therefore not q (invalid); and (d) modus tollens (if p then q; not q; therefore not p). The forms differed in that one involved items in which "at least part of one of the premises was contrary to observable fact"; the other involved concrete and familiar content. Results from the two were lumped together, even though the form with the concrete familiar items was easier. Criterion for "mastery" of a principle was 6 correct out of the 8 items for the principle.

Second Instrument: Four-card problems usually involve a given conditional rule and cards which may or may not be compatible with the rule. Here the rule was, "Whenever there is a number below the line, there is a letter above the line," and a letter, a numeral, an asterisk, or a masked region was in each of the top and bottom halves of the cards. Three four-card tasks made up the second instrument. These involved identification of correct domain for a masked top or bottom of a card, identification of card(s) incompatible with the rule, and identification of the half-masked cards which would test the rule. Criterion for mastery of a principle was correct decisions on all three cards which involved the principle. These tasks were administered by overhead projector displays, with answers recorded on a special form. This instrument was given right after the written test.

For each of the principles, the proportions of students meeting criterion on the two tests were compared within each grade and in toto.

4. Findings

(a) The proportion of students meeting criterion on the written test differed significantly from the proportion meeting criterion on the four-card instrument in all grade-by-principle combinations except for modus ponens with the twelfth graders ($p < 0.01$ and in most cases $p < 0.001$).

(b) For the valid principles, modus ponens and modus tollens, the written test was easier. For the invalid patterns, the four-card test was easier.

5. Interpretations

(a) The written form of an inference may lend itself to a formally learned "algorithmic" response, whereas a four-card task may call forth "native" logic.

(b) Concurrent validity for the two tests is not strong.

(c) Ordinary thought processes are not always consistent with mathematical logic.

Critical Commentary

(a) The results of this study provide an excellent illustration of the importance of a researcher's choice of instruments. As the investigator points out, content validity is no assurance of concurrent validity.

(b) One can always find (perhaps minor) points to object to. Here are four. First, are comments on reliabilities not appropriate for the sort of test used here? Second, the remark about the difference in performance on the two written versions was intriguing. Putting the results together makes one wonder how well-planned the analysis was. Some sort of correlational analysis would seem to have been in order. Third, should the written version have asked whether the conclusion was true? The contrary-to-fact items might then be particularly puzzling to students. Finally, it would appear that the criterion for "mastery" of a principle for the written-test version (6 out of 8 correct) was less stringent than the criterion on the four-card version (all 3 of the directly related instances correct). The investigator mentions the arbitrariness of the criteria. One wonders how greatly the results would have been affected if some different criteria had been used.

(c) The author is to be commended for not commenting on the apparent difference from grade to grade. After all, since only the stronger students may survive the natural selection of mathematics students from grade 8 to grade 12, the twelfth graders very likely were not comparable to the groups from the earlier grades.

(d) As always it is appalling to read the results of such status studies--can't our students, especially twelfth-year mathematics students, reason any better than that?! Jansson points out that few studies have examined the effect of instruction. We need such studies, with penetrating assessments of what we do teach. Can we teach more than logical algorithms? Can we indeed improve "native" logic, supported as it is with everyday non-mathematical uses of logical connectives?

(e) The author laudably examined several possible explanations of the results. In reading the discussion, however, one keeps wondering, "Why didn't he talk to any students?" Even researchers who look down on clinical studies acknowledge that interviewing students can provide potentially valuable data. How firmly convinced were students that their responses were correct? What sort of algorithms do the students use? (For example, mine "cancel" statements and combine what is left at the conclusion!) The approach to determining student thinking is to ask them to explain their responses; Adi, Karplus, and Lawson (1978) took such an approach and arrived at categories which seem to show a developmental trend.

SHOULD SCIENCE BE USED TO TEACH MATHEMATICAL SKILLS? Kren, Sandra R.; Huntsberger, John P. Journal of Research in Science Teaching, v14 n6, pp557-561, 1977.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by George W. Bright, Northern Illinois University.

1. Purpose

The purpose was to determine whether mathematical skills can be acquired by presenting either (1) quantitative science exercises alone or (2) science and mathematics exercises concurrently.

2. Rationale

A response to "demands to return to the basics" prompted an application of the "several theories of transfer of training" to science teaching in terms of the effects on mathematics achievement. Integration of the curriculum is presented as one way to achieve transfer. (It is not clear whether the references cited concerning attempts to integrate science and mathematics are reports of research or exhortations to integrate mathematics and science. The common element reported from these articles is the authors' beliefs that mathematics and science should be integrated.) The study was viewed as a follow-up to a study by Kolb (1968) in which greater achievement of science objectives was observed when a mathematics sequence preceded the science exercises.

3. Research Design and Procedure

The science exercises were from Science: A Process Approach I (prediction in various physical systems, and the measurement of angles). The mathematics instruction was lecture-demonstration based on a textbook supplemented by a packet of worksheets "to insure adequate coverage" (interpretation and construction of linear graphs; measurement and construction of angles).

The treatments were (1) science, (2) mathematics and science concurrently, (3) mathematics and (4) control (pre- and posttests only, with a 10-day lapse). The treatment lasted "approximately 12 consecutive school days."

Students were 161 fourth- and fifth-graders from eight classrooms. "Three classrooms were randomly assigned to treatment groups and five classrooms volunteered for one of the treatments used in the study " [emphasis added]. Distribution of classrooms among the treatments was not reported. The experimental design was a non-equivalent control group design.

Prior to the treatment, Sections I and II of the Kren Test were given. The test was administered again the fourth day after the treatment. Analysis of covariance, multiple range test, analysis of variance, multiple classification analysis, correlation, and item analysis frequency were used.

The Kren Test-Section I is on measuring and constructing angles; the Kren Test-Section II is on interpreting and constructing linear graphs. Each section is 15 items; the total reliability was .93. Content validity was endorsed by a panel of readers.

4. Findings

ANCOVA (the covariate is not clearly identified) was used on posttest scores to measure differences among groups. For each section of the test, significant differences ($p < .01$) are reported among the treatment groups. The df for each F-statistic is given as (3,157).

Mathematics-only and mathematics-science-concurrently were "equally effective" in teaching the material on angles, but science-only was "not an acceptable substitute." "Analysis of covariance was used to analyze the data in order to adjust for initial difference [among] the groups, which might have been present due to varying amounts of exposure to the prerequisite skills." The material on graphs was taught with equal effectiveness via mathematics-only, science-only, or mathematics-and-science-concurrently.

ANOVA was used on posttest scores to measure grade effect. A significant effect ($p < .01$) was reported only for Section I. Since means are not reported, it is impossible to tell which grade scored higher. Also, the df for the analysis is not reported.

For each section of the Kren Test, there was a positive correlation between pre- and posttest scores. There was also a positive correlation between Section I and Section II for the posttest only. Neither the correlation nor the significance levels were reported.

Constructing angles was harder at both grades than reading a protractor or measuring an angle. No statistics are reported for the multiple range test or the multiple classification analysis.

5. Interpretations

The science activity on prediction seems improperly placed. It should be delayed. Further investigations of the effectiveness of the integration of science and mathematics should be undertaken.

Critical Commentary

The study has obvious flaws. Of primary importance is that since classes were assigned to treatments (but not randomly) the unit of analysis should be the class mean. The correct df for the ANCOVA and ANOVA are (3,4) and (1,6) respectively. It is impossible to tell whether a reanalysis would result in any significant F-statistics.

The comments below are less critical.

1. The rationale for investigating the "integration of the curriculum" as part of the transfer of training paradigm is not clear. It is, of course, one aspect of the learning environment, but there is no justification that it is a critical aspect.

2. Twelve consecutive school days is long enough to anticipate some results, but the description is "approximately 12 days." Were the treatments of different lengths?

3. The nature of the use of the mathematics concepts in the science instructional treatment should be explained. This seems especially important if we are to learn to recognize the kinds of integration of curricula that are effective.

4. What did the control group do for 10 days? Why was the pretest/posttest lapse 10 days rather than 12?

5. What was the distribution of grades among the treatments?

6. The results are reported in a confused way. Some are in the discussion section but not the results section. Results for two analyses seem not to be reported at all.

7. The F-statistics alone are impossible to interpret. Means and multiple range test statistics are essential for the interpretation.

8. Differences in exposure to prerequisites among the treatment groups are cited as the cause for using ANCOVA to measure the treatment effect. It seems logical that there would also be differences in exposure to prerequisites between the grades. Why wasn't ANCOVA (instead of ANOVA) used to measure the grade effect?

As a whole, the study as reported lacks credence. The reader cannot determine from the information provided whether science can in fact be used to teach mathematical skills.

DESCRIPTION AND ASSESSMENT OF DIFFERENT METHODS OF TEACHING ENGINEERING STUDENTS MATHEMATICS. Macnab, D.; Mickasch, J. D.; Georgi, W. International Journal of Mathematics Education in Science and Technology, v8 n2, pp219-228, 1977.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Len Pikaart, Ohio University.

1. Purpose

The use of (1) tutorial sessions and (2) computer exercise sessions as part of the instructional procedures in a complex numbers course in a technical college are compared separately to the use of lectures only.

2. Rationale

Technical colleges or Fachhochschulen in West Germany offer four-year engineering degrees. The information explosion in the recent years has caused a general shift from instruction composed of "lectures, discussions, and exercises" to that consisting almost entirely of lectures. Using the results of a questionnaire to determine the priorities of mathematics topics taught in Fachhochschulen, the two alternative procedures were developed in earlier projects for integral calculus and complex numbers. Formative evaluations over a period of 1½ years involving 600 ~~students in 19 colleges indicated that tutorial groups were higher than~~ lecture groups in cognitive achievement and student satisfaction with the course and materials, but there was no significant differences between tutorial and computer exercise groups (no comparison is mentioned between computer exercise groups and lecture groups). Close examination of data in these earlier studies indicated that students who had "the lowest pre-instructional mathematics knowledge performed better in tutorial and computer groups than in the lecture group."

3. Research Design and Procedure

Two separate experiments are reported. In both, students enrolled in a complex numbers course were administered (1) an intelligence test, (2) a course entrance test (CET) designed to measure "pre-instructional mathematical knowledge", (3) a pre-test of 65 multiple-choice items, and (4) a post-test identical to the pre-test. A derived criterion measure was defined as:

$$\text{Performance } p = \frac{\text{Post-Pre}}{\text{Max-Pre}}$$

where Pre and Post are defined as the number of items correct on the pre-test and post-test respectively. Max is 65, the maximum number of correct responses in the pre/post-test. Study of this measure, p, indicates that it is a ratio of a gain score to the maximum possible gain for each student, which is expressed as a percentage in the report.

Students within a common CET group were assigned to one of three instructional methods: tutorial, computer, or lecture. Each group received ten instructional periods of 90 minutes. Six of the periods were tutorial sessions or computer sessions in the two associated groups and the remaining four periods were lectures. All ten periods were lectures for that group. The tutorial sessions were composed of about five students and conducted by another student who was a year more advanced. The computer sessions were designed for two students at a single interactive terminal with tutorial and drill-and-practice programs written in APL.

Both experiments used a treatment-by-levels design. In experiment 1, the treatments were tutorial group and lecture group whereas the levels were high, average, and low groups classified by CET scores. In experiment 2 the treatments were computer group and lecture group but the CET levels were high and low only, because the small number of terminals limited access to the computer. Twelve students were assigned to each treatment within each CET level. Thus in experiment 1 there were six treatment-by-level groups with a total of 72 students and in experiment 2 there were four treatment-by-level groups with a total of 48 students. The authors state that the mean for both treatment groups in a CET level were the same.

A Kruskal-Wallis one-way analysis of variance test was used for comparisons of levels within treatments and a Mann-Whitney U test was used for all pairwise comparisons. A level of .05 was employed for significance.

4. Findings

Table 1 is a listing of the mean performance index, p, as a percentage for the tutorial and lecture groups for the three groups of CET levels. An asterisk (*) indicates a pairwise significant difference, but note that the difference between means for high and low groups within the lecture treatment is also significant.

TABLE 1

EXPERIMENT 1: MEAN PERFORMANCE INDEX MEANS FOR LECTURE AND TUTORIAL BY CET GROUPS (N=72).

	C E T Groups			
	High		Average	Low
Lecture	64.8	*	45.9	43.6
			*	*
Tutorial	62.6		54.2	53.2

Table 2 presents similar data for experiment 2. Treatments are lecture and computer groups.

TABLE 2

EXPERIMENT 2: MEAN PERFORMANCE INDEX MEANS FOR LECTURE AND COMPUTER BY CET GROUPS (N=48).

	C E T Groups		
	High		Low
Lecture	58.3	*	41.3
			*
Computer	58.3	*	49.7

Spearman rank-order correlations between CET scores and performance scores for the various treatments follow:

Experiment 1		
Lecture	0.57*	
Tutorial	0.32	
Experiment 2		
Lecture	0.61*	
Computer	0.45*	*significant correlation

5. Interpretations

The authors interpret the results to indicate that the use of tutorial sessions and computer sessions were more effective for students with average or low CET scores than the use of lectures alone. For students with high CET scores, there appeared little difference in the use of the various treatments. The significance of the Spearman correlations for lecture groups was taken to indicate that performance was "highly dependent on pre-instructional behavior as measured by the course entrance test (CET)." This is also the case for the computer group but not for the tutorial group.

Critical Commentary

The major question to be raised in reading this report is, what effect did the definition of the performance index, p (see section 3, above), have upon the results of the study? This derived measure is non-standard and has potential to cause even more havoc than a simple change score. Associated with the use of the peculiar criterion measure is a question about why the investigators did not select readily available statistical models like analysis of covariance or a Lindquist Type I model (repeated measures). The conclusion of the authors about the superior effectiveness of tutorial and computer sessions as part of the course for students with average or low CET scores may be valid, but they are open to question until the criterion measure is defended.

COGNITIVE STYLE AND MATHEMATICS LEARNING: THE INTERACTION OF FIELD INDEPENDENCE AND INSTRUCTIONAL TREATMENT IN NUMERATION SYSTEMS. McLeod, D. B.; Carpenter, T. P.; McCornack, R. L.; Skvarcius, R. Journal for Research in Mathematics Education, v9 n3, pp163-174, May 1978.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Richard E. Mayer, University of California.

1. Purpose

The general purpose was to investigate aptitude treatment interactions (ATIs) in mathematics learning. In particular, the purpose was to determine whether there is an interaction between cognitive style and the amount of guidance given in instruction.

2. Rationale

Based on a careful review of the ATI literature, Cranbach and Snow (1977) and others have suggested that alternative instructional techniques should be tailored to the specific characteristics and aptitudes of individual students. In their choice of an aptitude, the authors used the dimension of "field dependence-independence" because it represents a "rather stable trait" that has been related to mathematics learning in previous studies (Witkin, Moore, Goodenough and Cox, 1977). In their choice of treatments, the authors, following Kilpatrick's (1975) suggestion, developed treatments relevant to mathematics education and also relevant to the theory of field-dependency--namely the amount of guidance and the presence or absence of concrete manipulatable objects. The main prediction is: students who score high in field independence should perform best under minimum guidance while students who score low in field independence should perform best under maximum guidance. An additional prediction is: cognitive style and amount of guidance may interact with the level of abstraction (i.e., presence or absence of concrete objects).

3. Research Design and Procedure

The subjects were 116 prospective elementary school teachers, with 81% being women. There were four treatment groups: (1) Min-M, minimum guidance with manipulatives, (2) Min-S, minimum guidance with symbols only and no manipulatives, (3) Max-M, maximum guidance with manipulatives, (4) Max-S, maximum guidance with symbols only. All subjects learned the same information, namely how to add and subtract in base four and base five under one of these four treatments. The manipulatives were multi-base arithmetic blocks. Four dependent measures for each subject were taken following learning: (1) QS-Posttest: addition, subtraction, multiplication and division problems in base three, in which the subject was not allowed to use the blocks; (2) QM-Posttest:

addition, subtraction, multiplication and division problems in base six, in which subjects were taught how and encouraged to use the blocks; (3) QS-Retention: same as above given four weeks later; (4) QM-Retention, same as above given four weeks later. The aptitude test was a version of the Hidden Figures Test (HFT) and a pretest was also given that tested for general knowledge about non-base ten number systems.

4. Findings

The regression coefficient for each treatment group was determined by relating score on the Hidden Figures Test (HFT) to adjusted score on the dependent measure for each of the four dependent measures. For QS-Posttest, the regression coefficients were .05, -.07, .14, and -.05 for Min-M, Max-M, Min-S and Max-S, respectively; for QM-Posttest, the corresponding coefficients were .14, .02, .03, and -.23; for QS-Retention, .12, .03, .22, and -.08; for QM-Retention, .10, .07, .09, and -.05. Tests for interactions between amount of guidance and score on the HFT were significant for QM-Posttest ($p < .006$) and QS-Retention ($p < .029$), marginally significant for QS-Posttest ($p < .092$), and not significant for QM-Retention ($p < .222$). The only significant interaction involving abstractness and score on the HFT occurred for QM-Posttest ($p < .012$).

5. Interpretations

These results provide clear support for Witkin's hypothesis that field independent subjects should perform better if they are allowed to work independently while field dependent subjects should perform better if they are given high levels of guidance. There were significant ATIs for two of the four dependent measures, and the appropriate trend was present in the other two. In general, the performance of maximum guidance subjects was negatively related to how high they scored in field independence while the performance of minimum guidance subjects was positively related to how high they scored in field independence. Apparently, level of abstraction was not an important factor in this study. The results have implications for mathematics instruction; for example, the fact that most mathematics textbooks provide high guidance (and no manipulatives) suggests that they may be less effective for high field-independence students.

Critical Commentary

This study is a case example of how to perform a good ATI study. The authors shunned the "shotgun" approach of throwing in many possible aptitude measures and many treatments and seeing what ATIs come out. Rather they carefully chose an aptitude (field dependence) and a treatment (amount of guidance) that were theoretically related to one another and to their task (mathematics). They offered a priori predictions based on established theories of cognitive style and discovery. They tested the theories in a well-designed, clear study that was directly related to mathematics instruction. Finally, they analyzed and discussed the nature and importance of the ATIs they obtained.

Because the authors have based their work on interesting theories rather than dealing with ATIs on a purely empiricist level, their findings have general implications for advancing both theory and instruction. The results are significant because they confirm Cronbach & Snow's (1977) contention that ATIs exist, and that they can be uncovered by careful theory-based research. Further, the results provide an independent line of support for the existence of the dimension of field dependency, and point to its relation to discovery instruction. The authors do not provide a detailed discussion of the implications for mathematics education; however, in general ATIs are interpreted to mean that one method of instruction should be used for some students while another method should be used for others. If discovery methods result in broad learning, then an alternative strategy would be to provide field-dependent learners with the needed tools to succeed at discovery; for example, there is some suggestion in this study that concrete manipulatable objects may be an aid in this case. Recent work on ATIs in mathematics by Egan & Greeno (1974), involving discovery makes a similar point.

There are several additional, minor points that could be improved in this paper. First, the reason for using "levels of abstraction" is not adequately presented, nor are the predictions clearly justified. The discussion of the results involving this variable is also weak. Second, the data are heavily adjusted and analyzed; that is certainly not a criticism, but it would also be useful for the reader to see a summary of the raw data. In addition to the useful regression techniques, an alternative is to partition each treatment group into field independent, field dependent and neutral (based on HFT scores); then the average posttest scores could be given for each of these three subgroups for each treatment group.

Reference

Egan, Dennis and Greeno, James G. Theory of Rule Induction: Knowledge Acquired in Concept Learning and Problem Solving. Knowledge and Cognition. Potomac, Maryland: Gregg, 1974.

THE EFFECTS OF INSTRUCTION IN SENTENTIAL LOGIC ON SELECTED ABILITIES OF SECOND- AND THIRD-GRADE CHILDREN. McGinty, Robert. Journal for Research in Mathematics Education, v8 n2, pp88-92, March 1977.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Helen Adi, Northern Illinois University.

1. Purpose

"The main purpose of the study was to determine the effect of different types of instruction on second- and third-graders' ability to derive valid logical conclusions from verbally expressed hypotheses... The effects of the training were also compared in terms of the subject's performance on perceptual reasoning and classification tasks."

2. Rationale

The development of logical reasoning abilities in elementary school children is of concern to mathematics educators. Children of ages 6 to 8 recognize valid logical conclusions derived from verbal premises (Hill, 1961), but fewer children are able to test the logical necessity of a conclusion (O'Brien & Shaprio, 1968). Further research had also indicated that elementary school children performed better on sentential logic when given specific instruction in logic.

How do different instructional treatments in logic compare in terms of second- and third-graders' performance on three different posttests of sentential logic, perceptual reasoning, and classification tasks? An attempt to answer such a question provided a pragmatic rationale for conducting the study. No theoretically-based argument for choosing the specific instructional treatments or the dependent measure variables was presented.

3. Research Design and Procedure

A sample of 16 classes of second- and third-grade students, eight at each grade level, was selected for the study. Each class consisted of approximately 25 students.

Four instructional treatments in logic were administered to independent groups of students. Two classes of each grade level were randomly assigned to each treatment. One control group received no instruction in sentential logic, while the other three experimental groups received instruction in logic using different sets of materials. The same operations of intersection, union, conjunction, disjunction, negation, and rules of inference were introduced in all three experimental treatments.

Sentential logic, perceptual reasoning, and classification defined.

the dependent variables. Three posttest scores were obtained for each measure. Sentential logic was measured by a 30-item test. The items could be correctly answered with "yes", "no", or "maybe". Perceptual reasoning was measured by a 30-item test adapted from Raven's Colored Progressive Matrices Test (1959). Classification was measured by another 30-item test from Raven's Classification Test (1970).

A repeated measures design was used to analyze the data. Grades and treatments were fixed factors, and test administrations defined the repeated measure factor. The class mean was the unit of analysis. Thus, a total sample size of 16 was equally distributed among eight cells.

4. Findings

(a) For sentential logic, there were significant main effects ($p < .01$) of grade and treatment, and a significant interaction effect of grade x treatment.

(b) For perceptual reasoning, there were significant main effects ($p < .01$) of grade and test administration.

(c) For classification, there were significant main effects ($p < .01$) of grade, treatment, and test administration, and a significant interaction effect ($p \leq .01$) of grade x text administration.

5. Interpretations

The interpretations of the study were:

(a) "Second and third grade children can have some success in answering certain items from sentential logic when they are exposed to selected instructional materials."

(b) "Students retained the positive effects of instruction over a period of time" (three weeks).

(c) "Practice on items in sentential logic did not seem to increase test scores, although practice on items in perceptual reasoning and classification did seem to increase test scores."

Critical Commentary

Three different "types of instruction" in sentential logic were selected, and their effects on sentential logic, perceptual reasoning, and classification were compared. The three "types of instruction" used Furth's (1970) logic materials, Dienes and Golding's (1966) logic materials, and set theory to explain the logic operations and principles. However, no explicit conceptual reason was provided by the author to explain the selection of these "types of instruction". Along what variables did the treatments differ?. And how were these variables controlled? Is the fact that some of these treatments proved to "work out" with younger children in previous research a sufficient theoretical reason for conducting another study? Did the author intend to replicate previous studies?

The study "attempted to expand on (sic) previous results by comparing the performance of subjects assigned to three different treatments on three different posttests" of sentential logic, perceptual reasoning, and classification. First, why were the selected abilities of sentential logic, perceptual reasoning, and classification specifically chosen? And second, why did the author expect the treatments to have differential effects on these selected abilities? It was left to the readers to provide possible answers to such questions.

Sentential logic abilities of second- and third-grade children were measured by their performance on a 30-item test where the items could be correctly answered by "yes", "no", or "maybe". Performance on such a test does not represent the ability of a person to reason logically. Recognition of valid conclusions is a different capability from generating these conclusions. A test-retest reliability of .69 simply means that the students may have been consistent in their responses to some degree. Two students may consistently and correctly choose a correct "yes" response for different logical justifications. Such a hypothesis is confirmed by recent research on intellectual development.

The design of the study called for analyses of variance with a repeated measures design. Classes were randomly assigned to treatments, and thus class means were correctly considered as the units of analysis. However, this reduced the sample size to 16, with n = 2 in each cell. The application of appropriate non-parametric tests may have been more suitable.

What is the contribution of the present work to the body of knowledge on the teaching and development of logic in young children? Although the findings were not qualitatively different from previous results in the literature, this study may be considered by some as another confirming evidence that teaching sentential logic to second graders is possible, and effective in terms of such selected abilities of children as classification. Don't we expect our first graders to solve multiple classification tasks correctly? Do we also have to teach our first graders sentential logic?

THE RELATIONSHIP BETWEEN THE MATHEMATICAL STRUCTURE OF EUCLIDEAN TRANSFORMATION AND THE SPONTANEOUSLY DEVELOPED COGNITIVE STRUCTURES OF YOUNG CHILDREN. Moyer, John C. Journal for Research in Mathematics Education, v9 n2, pp83-92, March 1978.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Jane Swafford, Northern Michigan University.

1. Purpose

The study sought to assess the compatibility of the mathematical structures of Euclidean transformation concepts and the "spontaneously developed" cognitive structures of children ages 4 to 8. More specifically, the study addressed three questions:

a) Are children's understanding of translations, reflections, and rotations dependent on the presence of explicit physical motions?

b) Are reflections easier for children than translations and rotations?

c) Do children progress from using the topological relation of "surrounded by red" to using relatively more complex projective and Euclidean relations?

2. Rationale

Since instructional programs are commonly based on the mathematical structure of a concept, a delineation of the relationship between cognitive and mathematical structures would seem to have profound implications for the development of effective curricula.

Because isometries are also homeomorphisms and translations and rotation are compositions of reflections, topological relations and reflections can be considered as mathematically primitive. Although Piaget has observed that topological concepts develop first cognitively, previous research has found that translation is easiest for children. Further, translation is, mathematically a static one-to-one correspondence, while cognitively children may need to see motion.

3. Research Design and Procedure

Twenty-four children selected randomly from each of grades preschool through third (Total-120) were tested. The testing materials consisted of pairs of transparent plastic circles. In one set, the circles were half red and half clear with a black diameter. Another set was clear except for a black diameter with a thicker half-diameter.

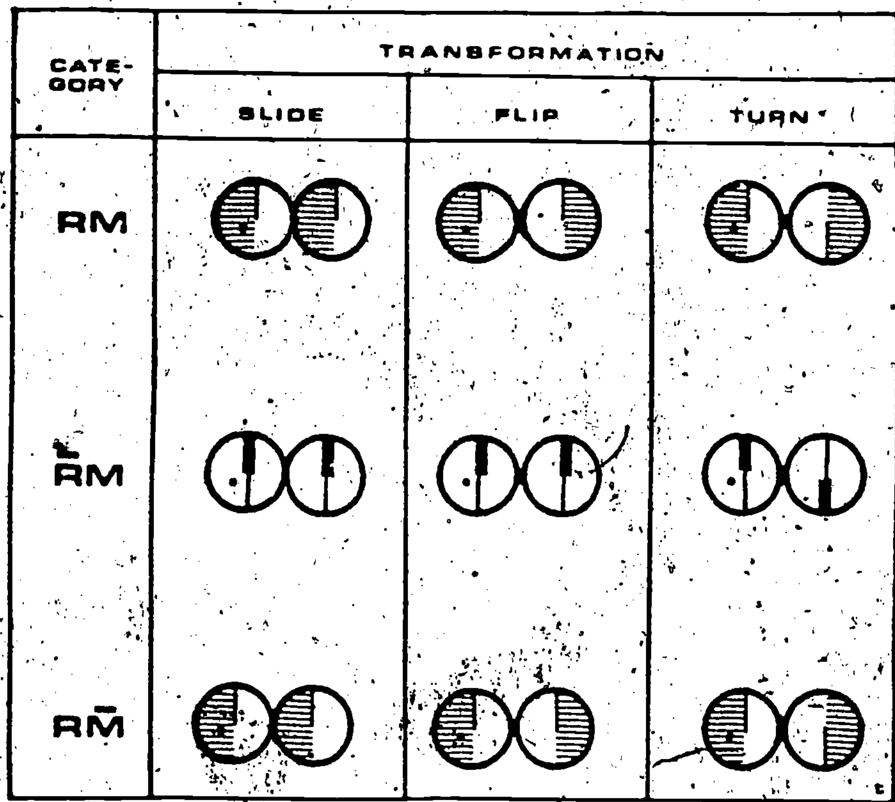


Figure 1. Position of the circles for the nine tasks.

(See Figure 1.) Nine tasks were presented to each child. The tasks correspond to the three Euclidean transformations under each of three conditions: red circles with motion (RM), clear circles with motion ($\bar{R}M$), and red circles without motion ($R\bar{M}$). Within grade levels, children were randomly assigned to different task sequences. Instructions for each task were presented by cassette tape and earphones.

In each task, tangent circles were placed before the child. In the motion tasks, the appropriate transformation was demonstrated by the experimenter. The child was asked to draw a dot on the right-hand circle to correspond to the one made by the experimenter on the left-hand circle. In tasks without motion, no motion was demonstrated nor "was any motion made on tape."

4. Findings

Children were scored on each of the nine tasks as follows:

0 = Dot in wrong half
1 = Dot in correct half, wrong quadrant
2 = Dot in correct quadrant, outside of tolerance limits
3 = Dot within tolerance limits

Relative performance on pairs of tasks were compared. Subjects were classified as doing poorer, the same, or better on one task than another. In most cases, grade levels were combined. Chi-square statistics for 1x3 contingency tables were used to analyze the independence of these three classifications of students (poorer, same, better) and the variable of interest.

Comparisons of the RM and $R\bar{M}$ tasks for each of the three transformations indicate that only for rotations is performance dependent on explicit demonstration of the motion. Comparisons of pairs of transformations under each of the conditions RM, $R\bar{M}$ and $\bar{R}M$ indicate that translations are no more difficult than reflections while rotations are the most difficult. Comparisons of RM and $\bar{R}M$ tasks for each of the transformations indicate that the effect of being surrounded by red does not decrease with age.

Examination of the distribution of the scores by grade level shows the frequency of higher scores increasing with grade level, indicating the use of more complex projective and Euclidean relations with age.

It was noted also that grade level was related to performance but IQ was not, although no data were cited.

5. Interpretations

With warnings that these results should be interpreted cautiously, the author concluded that mathematical and cognitive structures do not always agree. In some cases, mathematically primitive notions (topological relations) precede more complex ones. In other cases, children are no more successful with mathematically primitive notions (reflections) than more complex ones (translations). Hence, programs based on mathematical structures can only be considered as a starting point for curriculum development. The author suggested that further research might indicate that the emphasis on motions is misplaced. Furthermore, children might not classify transformations into the distinct mathematical categories.

Critical Commentary

The study was carefully designed and executed. Whether it represents more than an exercise in research methodology is debatable. Both its theoretical rationale and interpretations are suspect.

Theoretically, there is a sense in which both topological relations and reflections may be considered primitive. There is also a sense in which they may be considered more sophisticated. Furthermore, the assumption that curriculum is often built on primitive mathematical structures can be challenged with counterexamples. In addition, since controls were not exercised on the learning environment and since grade levels were combined in most cases, neither the "spontaneity" nor development" of the cognitive processes was in reality investigated.

Discounting the theoretical rationale, the study can still be viewed as a very careful, but restricted, investigation of children's understanding of Euclidean transformations. At the level of the three specific questions raised, the study is of some interest. However, whatever its technical merits, this study cannot be construed to have enlightened the understanding of the relationship between mathematical and cognitive structures nor to have any serious implication for curriculum development.

It should be noted that the two circles used in each task were tangent. The flip task appears to be a single figure with line symmetry. The slide task did not have far to slide. Under these circumstances, even the hint that the current curriculum "emphasis on the motion aspect of motion geometry should be made advisedly" seems presumptuous. Also, it should be no surprise that the flip was at best only slightly harder than the slide. In fact, results of the RM tasks would suggest that only with red cues were the slide and flip task comparable.

Finally, the expected values for each cell in the 1x3 contingency tables were calculated on the assumption that each is equally likely. But if each score (0,1,2,3,) is equally likely, then for any individual P (task A score < task B score) = P (A < B) = $\frac{12}{16}$ while P (task A score = B score) = $\frac{4}{16}$. This correction does not modify the significance of the Chi squares. However, it does affect the interpretation of the sources(s) of the significance. Consequently, even the interpretations of the more restrictive study must be made with caution.

THE DIFFERENCE IN LEVEL OF ANXIETY IN UNDERGRADUATE MATHEMATICS AND NON-MATHEMATICS MAJORS. Ohlson, E. LaMonte; Mein, Lillian. Journal for Research in Mathematics Education, v8 n1, pp48-56, January 1977.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Marilyn N. Suydam, The Ohio State University.

1. Purpose

The study attempts "to determine whether a difference existed in the degree of anxiety possessed by undergraduate mathematics majors as compared to undergraduate nonmathematics majors. A secondary purpose was to investigate the extent to which difference in sex influences the degree of anxiety" of these two groups.

2. Rationale

Previous studies dealing with anxiety have explored the need to recognize the existence of anxiety in the classroom, anxiety in relation to testing conditions, the effect of anxiety and other emotional factors on learning, and the relationship of anxiety to motivational clues. "Although researchers seem to be in agreement that anxiety does affect learning, there is a lack of agreement among researchers as to the etiology of anxiety. Some studies have gone as far as to explore the possible effects of anxiety on such nonbehavioral phenomena as college grade point averages, intelligence, and task complexity." But "there is a lack of research relating anxiety level to various academic disciplines, particularly mathematics. Those studies dealing with mathematics have been mainly concerned with attitudes. The present study, therefore," was conducted for the purposes stated above.

3. Research Design and Procedure

The population consisted of the 11,000 undergraduate students enrolled spring quarter 1973 at the University of Northern Colorado. From this population, 124 nonmathematics majors (80 female, 44 male) and 67 mathematics majors (34 female, 33 male) were chosen randomly. "Both groups were comparable in terms of 'major' requirements; because a regression model was used to analyze the data, the proportionality assumption did not have to be met. Therefore, maintaining an equal number of students at each academic level, major, or sex was not necessary."

Anxiety, achievement, and aptitude were used as predictors "since the literature indicated that a relationship exists" between them. Sex was used as a predictor "since the literature review indicated both sexes need to be used in anxiety studies." The predictor variables follow.

Y	anxiety score from A-State subscale of STAI
X_2	anxiety score from A-Trait subscale of STAI
X_3	mathematics major (1 means yes, 0 means no)
X_4	academic level (1 means freshman; 2, sophomore; etc.)
X_5	sex (1 means female, 0 means male)
X_6	GPA from winter quarter 1973 (if student teaching that quarter, fall quarter 1972 was used)
X_7	cumulative GPA
X_8	ACT mathematics standard score
X_9	ACT composite standard score
X_{10}	classroom membership (1 means a student was in a mathematics class, 0 means a student was not in a mathematics class)
X_{11}	$X_3 \times X_4$

Both subscales (A-State, A-Trait) of the State-Trait Anxiety Inventory (STAI) were administered to each S in "a classroom atmosphere." Ward's multiple linear regression model was used to determine the contribution of sets of the eleven predictor variables to the variability of the criterion, A-State anxiety. Factor analysis (principal axis method, then varimax) was used to determine the groupings of the predictors. Finally, Restricted Models (the Full Model minus one factor) were each tested against the Full Model, to determine the unique contribution of each factor to the system. "A large drop in RSQ would indicate that the variables in the RM were making a unique contribution to the predictive efficiency of the criterion, A-State anxiety. A small or zero drop would indicate that the set is not adding anything independently of the other variables. If the drop is not significant, then further testing of subsets of those variables is unnecessary. This is one reason for the hierarchical grouping."

4. Findings

Five factors were determined:

(a) Mathematics variables: mathematics major (3), class membership (10), mathematics major x academic level (11)

(b) Achievement variables: sex (5), GPA last quarter (6), cumulative GPA (7)

(c) Aptitude variables: ACT mathematics (8), ACT composite (9)

(d) General anxiety variable: A-Trait score (2)

(e) Academic level (4)

Intercorrelations were provided in Table 3, varimax rotation loadings are in Table 4, and, from Tables 2 and 4, the predictor loadings on factors were presented in Table 5. Table 6 shows the "schematic for regression models"; the RSQ for the Full Model (FM) compared with the criterion (A-State score) was .4692. Results of testing each factor were:

FM minus Factor 1	.4575	
FM minus Factor 2	.4640	
FM minus Factor 3	.4677	
FM minus Factor 4	.0567	- significant drop (a drop of .05 is considered significant)
FM minus Factor 5	.4688	

5. Interpretations

"The resulting statistical analysis led the investigators to conclude that:

(a) mathematics majors are not more anxious than nonmathematics majors as measured by the STAI

(b) anxiety levels of mathematics majors did not increase as their academic level increased from the freshman class through the senior level

(c) being in a mathematics classroom created no more anxiety than being in a nonmathematics classroom

(d) the sex, current GPA, and cumulative GPA variables did not contribute to the prediction of A-State anxiety

(e) general anxiety scores (A-Trait) can be used to predict specific anxiety scores (A-State)."

In discussing these various conclusions the authors state:

(a) "Perhaps the fact that a student is a mathematics major implies that he has had some success in the area and, therefore, felt no more anxious in a mathematics class than in any other class. . . . But what about the anxiety level of nonmathematics majors in mathematics classes? . . . there appeared to be no difference in anxiety levels for different classroom situations. . . . Another explanation might be that stressful conditions are necessary to produce significant anxiety levels in mathematics majors. . . . Without evaluation, one classroom situation may be just like any other classroom atmosphere and, therefore, no measurable difference in the anxiety levels of students exist. Before we can really conclude that mathematics does not create anxiety in students, we need to explore these other possible suggestions."

(b) "Perhaps for sex differences to be noted in an anxiety study, there must be imposed threat of some kind . . . involving pain and physical danger."

(c) "One explanation for finding that GPAs of students are apparently independent of their anxiety levels might be (that) 'perhaps for attitudes to interact with achievement they have to be extreme, and extreme attitudes . . . may be rarer than is commonly thought'."

(d) ". . . the students are quite successful in their achievement. This could imply that successful students are nonanxious students. . . . Evidently these students have enough confidence to do well regardless of their anxiety level."

The authors summarize the report by querying whether dislike, fear, or anxiety are directed toward mathematics or toward the classroom atmosphere. Of the variables tested, only general anxiety had any significant predictive efficiency. More research is needed.

Critical Commentary

This report might be titled, "The Report That Leaves One Wondering, If Not Anxious." It leaves one wondering what the words mean, what the purpose was, what might be missing, and how it came to be accepted by reviewers for publication in a reputable journal. One even suspects a hoax. Three other reviewers refused to review the report. One said:

> "I tried to abstract the article. However, it is so poorly written, designed, and thought through, I have decided it is not worth my time to review or I.M.E.'s share to include. It is research gobbledy-gook at its worst. How did it get in JRME?"

Another reviewer said:

> "I have to agree (with the two previous reviewers)--this may be the most confused article I have seen in JRME. It *is* pure gobbledegook! I do not see how it survived review!"

There were moments when I wondered whether the principle that all articles could be reviewed for I.M.E. was really valid!

The review of previous literature may be the clearest portion of the report. The authors cite studies by psychologists, but are unaware of any previous research on anxiety in mathematics (although they do cite a study later). They noted Aiken's review on attitudes, but apparently failed to locate Aiken (1970) or Aiken (1976).

When one discards unwarranted jargon, reorganizes paragraphs, and ignores nonsensical statements (or, at least, statements which do not seem to fit into the context), one finally realizes that this is a study of general anxiety, not anxiety toward mathematics. The sample consists of mathematics majors and nonmathematics majors who are either taking or not taking a mathematics course. The same sample could have been divided into English majors and non-English majors who are either taking or not taking an English course--or a mathematics course. But knowing that doesn't resolve some of the questions that occur about the sample, such as why they selected a sample with such a disparity in the number of males and females in the two groups (the nonmathematics groups is predominantly female), or why there was a disparity in the size of the two groups in the first place (the nonmathematics group is almost twice as large as the mathematics group). One suspects that, after randomly selecting the sample, they looked to see what they had, but this is not stated.

The STAI was administered in "a classroom atmosphere." What classroom? Under what conditions? Were those taking mathematics in a mathematics classroom? It would seem imperative to know the conditions under which an anxiety measure was administered. "Classroom atmosphere" conveys virtually nothing.

One really wonders if two articles became mixed up--pages 50-52, explaining the factor analysis and regression analysis, seem to come from another manuscript. The authors test for anxiety, report a factor analysis and perform some subtractions of variable effects, and conclude that one group of subjects differs from another. How they made the leap is totally unclear. No analysis is reported of data comparing mathematics and nonmathematics majors; there is insufficient information given to guide the reader to make the leap. Many things are possible with regression analysis--but readers are rightfully very wary of believing that regression resolves all questions.

Credibility is further weakened by the discussion, where some of the statements seem to be plain nonsense. They have assembled a random collection of quotations about programmed materials without teacher threat, evaluation, and so on, and combined them with their own biases. Moreover, the researchers confuse anxiety and attitudes in the discussion: they seem unaware that the two constructs differ--and that the STAI was not assessing attitudes per se.

Given the explanation in the report, their conclusions cannot be accepted: the researchers failed to communicate. Anxiety about mathematics will not go away by saying that there is no anxiety present.

References

Aiken, Lewis R., Jr. Nonintellective Variables and Mathematics Achievement: Directions for Research. Journal of School Psychology 8: 28-36; March 1970.

Aiken, Lewis R., Jr. Update on Attitudes and Other Affective Variables in Learning Mathematics. Review of Educational Research 46: 293-311; Spring 1976.

INTERACTIVE EFFECTS OF PRIOR MATHEMATICS PREPARATION AND LEVEL OF INSTRUCTIONAL SUPPORT IN COLLEGE CALCULUS. Pascarella, Ernest T. American Educational Research Journal, v15 n2, pp275-285, Spring 1978. (a)

STUDENT MOTIVATION AS A DIFFERENTIAL PREDICTOR OF COURSE OUTCOMES IN PERSONALIZED SYSTEM OF INSTRUCTION AND CONVENTIONAL INSTRUCTIONAL METHODS. Pascarella, Ernest T. Journal of Educational Research, v71 n1, pp21-26, January 1977. (b)

INTERACTION OF MOTIVATION, MATHEMATICS PREPARATION, AND INSTRUCTIONAL METHOD IN A PSI AND CONVENTIONALLY TAUGHT CALCULUS COURSE. Pascarella, Ernest T. Audio Visual Communication Review, v25 n1, pp25-41. (c)

Expanded Abstract and Analysis Prepared Especially for I.M.E. by James W. Wilson, University of Georgia.

1. Purpose

The purpose in all three studies was to compare Personalized System of Instruction (PSI) and conventional (lecture) instruction in an introductory calculus course, and to search for interactions of method with prior mathematics preparation and/or motivation.

2. Rationale

The PSI is a highly structured, self-paced program in this Syracuse University calculus course. It was hypothesized that students with relatively low mathematical preparation would achieve better with the high instructional support condition of PSI than with the low instructional support lecture condition whereas highly prepared mathematics students would achieve equally well under the two instructional conditions. Similarly the self-posed PSI should benefit the highly motivated student over the less motivated, whereas the lecture would not.

3. Research Design and Procedure

All three studies used students from the first calculus course of a four-course sequence at Syracuse University. In study (a) there were 60 PSI students and 188 lecture students; studies (b) and (c) each used 47 PSI students and 47 lecture students randomly selected from the respective groups.

A pre-post non-equivalent control group design was used in each study and the same set of variables was measured: a mathematics placement examination to measure previous mathematics preparation, the Stern Activities Index personality inventory (including a measure of motivation), and a 132-point, eight-question end-of-semester examination. Study (b) also reported an attitude measure for an additional dependent variable.

Studies (a) and (b) were analyzed using multiple regression analysis; study (c) was analyzed with a 3 (levels of mathematics preparation) X 3 (levels of motivation) X 2 (instructional treatments) factorial design using a least squares analysis of variance. Regions of non-significance were determined in (a) and (b) by the Johnson-Neyman technique.

4. Findings

An interaction of level of preparation and instructional method was found in (a) and (c). In each case the PSI instruction produced achievement at about the same level for low, medium, or high prior mathematics preparation.

An interaction of level of motivation and instructional motivation was found in (b) and (c). Students with high motivation in PSI instruction tended to perform better than those with low motivation level. Under the lecture condition, motivation level was unrelated to achievement. A similar interaction was found for the attitude measure in (b).

There was a main effect for instructional method in all three studies with the PSI method leading to higher achievement.

5. Interpretations

The higher performance with PSI methods should be interpreted in light of the interaction effects. That is, students who benefit most from the PSI instruction may have certain aptitudes (relatively low prior mathematics preparation) or traits (high motivation).

Critical Commentary

These studies are technically well-done and carefully reported. The author is very cautious in discussing the results and candidly remarks on the limitations of the design. For example, the non-experimental design is a serious weakness, but it was dictated by the available instructional setting and the author provides information directed to the sources of invalidity of such a design.

The reports failed to present any analysis of the range of content covered under either treatment. The PSI students were self-paced and could have covered more material than the lecture group. Some analysis of the coverage of material in the treatments <u>and</u> in the eight-question 132-point examination is essential for understanding the results.

None of these articles indicates a date for the study. The mathematics placement examination was available in 1972 and all three studies were reported by early 1977 (study (a) was presented at the 1977 AERA Annual Meeting). There is no reference to the other two studies in any of the reports. Hence, what we have is not a program of studies, one building on the other, nor replications, but rather the same study repeated, or reported, three times.

Finally, the problem is of interest primarily for the understanding and evaluation of the particular application of PSI to calculus instruction at Syracuse. The results may not generalize to other applications of PSI and to other contrasts of methods.

TEACHERS', PRINCIPALS', AND UNIVERSITY FACULTIES' VIEWS OF MATHEMATICS LEARNING AND INSTRUCTION AS MEASURED BY A MATHEMATICS INVENTORY. Post, Thomas R.; Ward, William H. Jr.; Willson, Victor L. Journal for Research in Mathematics Education, v8 n5, pp332-344, November 1977.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Marilyn J. Zweng, University of Iowa.

1. Purpose

The intent of this study was to determine if mathematics teachers, secondary school principals, and university mathematics educators have the same or dissimilar views about methods of teaching mathematics, goals of mathematics instruction, and other educational issues specific to mathematics learning and teaching. The attitudes of all three groups were assessed by the same 30-item questionnaire.

2. Rationale

In a prior examination of the data for teachers and principals (the same data utilized in this study exclusive of the data from university educators), the 30 items of the questionnaire had been sorted into seven factors: (1) Flexibility, (2) Mathematics as a Process, (3) Teacher Concern for Student, (4) Vocational Satisfaction, (5) Nonrigid Practices, (6) Attitude Toward Teaching, and (7) Higher Order Concerns. The analyses suggested principal-teacher differences for two of the factors, Teacher Concern for Student and Higher Order Concerns. Because university mathematics educators are a population whose attitudes toward mathematics instruction are also of interest, the original intent of this study was to sample their responses to the same questionnaire and compare their attitudes with respect to the seven factors to those of the other two groups. It was not possible to sort the items into the seven subsets when the university educators' responses were incorporated in the data, so overall and item-by-item comparisons only were examined.

3. Research Design and Procedure

The Questionnaire. Principals (p), mathematics teachers (t), and college and university level mathematics educators (u) were asked to respond to the same 30-item inventory entitled, "The Mathematics Inventory for Teacher (MIT)." (The inventory is included as an appendix to the journal article.) Each item of the inventory is a statement about some aspect of mathematics teaching, the nature of mathematics, the goals of mathematics instruction, or some other professional attitude. For example, in Item 1 the subject is asked to react to the statement, "The field of mathematics consists primarily of procedures and formulas which are used in many occupations and in everyday experiences." Item 10 states "I regularly (at least once a week) inform my students of the learning progress they are making by giving assignments, quizzes, or tests." A four-point scale was provided for response to each item: (1) strongly agree, (2) agree, (3) disagree, (4) strongly disagree. University educators and principals were asked to respond to the MIT "as you would expect the ideal mathematics teacher to answer the items."

The Sample. Complete MIT data were available from 160 principals, 199 teachers, and 117 college and university level mathematics educators. The secondary school principals were randomly selected from California, Michigan, and Indiana. In all, 222 principals were asked to participate. These principals were then asked to select randomly one mathematics teacher from their faculty for the study. A total of 200 university level mathematics educators were randomly selected, by geographic region, from the mailing list of the NCTM's Bulletin for Leaders. Since most teachers and principals occurred in the data as matched pairs, the teacher and principal responses could not be assumed to be independent. However, very low correlations were found to exist between the two groups' responses for almost all of the 30 items, so it was felt that teachers and principals could be treated as if they were independent groups.

4. Findings

Multivariate analysis of variance procedures were used to test quality of the means among the three groups overall for the 30 items, and each pairwise contrast ($\bar{p} - \bar{t}$, $\bar{p} - \bar{u}$, $\bar{t} - \bar{u}$). Differences were found to exist overall and between each of the three pairs with $p < .01$ in each instance. An item-by-item analysis revealed that a difference between principals' and teachers' means existed for 10 out of the 30 items, a difference between principals' and university educators' attitudes occurred for 23 items, and the differences between the means of teachers' and university educators' scores were significant for 18 of the 30 items. There was "agreement" among the three groups on only five items.

5. Interpretations

The analyses of the data revealed large-scale discrepancies among the three groups. The greatest divergence in attitudes about mathematics instruction occurred between principals and university educators. Second in order was the difference in attitude between mathematics teachers and university educators. The smallest discrepancy was between teachers and principals, but these two groups, nevertheless, exhibited differences of opinions on one-third of the items. An extensive discussion of these differences and possible explanations are provided in the article.

Critical Commentary

The results of this study are certainly not surprising. They merely substantiate the widespread belief that "ivory tower" university educators hold idealistic views which are far-removed from the real world of the classroom. Since the results agree with commonly held assumptions, this would probably have been the sum and substance of this abstractor's comments-- if the study had been reported as most are-- with tables of F-ratios, t-ratios, and correlations, but no raw data. However, the authors, most commendably, included an appendix which contained all 30 items on the inventory and the mean scores for each of the three groups for each item-- a wealth of information, and a tremendous temptation for one who views "playing with data" as a leisure-time game.

It appears to the abstractor that the variable which the investigators studied was intensity of attitude since the values of the mean scores (the focus of the analysis) indicate whether a group agreed slightly or strongly, or disagreed slightly or strongly with a position. An examination of the data from another perspective suggested that it might be interesting to look, instead, at priorities of attitudes. For example, the researchers' analysis of the mean scores for Item 8 (principals-- 1.35, teachers-- 1.38 and university educators-- 1.17) showed a significant difference between university educators and each of the other two groups for this item, but all three groups gave Item 8 their lowest score (thus agreed most strongly with this item). In terms of ranking, each group had placed this item in identically the same position. Could it be that priorities are the same for all three groups, but one group simply takes a stronger position? To answer this question, the 30 items were rank ordered from "most strongly agree" to "most strongly disagree" for all three groups, with the results shown in Table 1. Table 1 lends support to the contention that the three groups varied in the strength of their responses. The first observation is that the range of scores for mathematics educators (1.17 to 3.48) is considerably greater than the range of scores for the other two groups. The neutral point, which is 2.5 on the MIT's four-point scale, is also of interest. Examining the interval (2.25,2.75), it can be observed that for principals, there are six MIT items having means scores in this range; for teachers, eight items had mean scores in the range; but for mathematics educators, only two items have mean scores between 2.25 and 2.75. Together, these observations tend to suggest that university educators take a more polarized position than principals and teacher, but they do not necessarily imply that university educators have different priorities.

In order to answer the question about priorities of the three groups, Table 2 was developed in which the ranking of each item for each of the three groups is displayed. Tied ranks were assigned the average rank which the two scores occupied, as is conventional practice. Spearman's Coefficient of Rank Correlation, ρ, was computed for each pair of comparisons. The coefficient of correlation between the principals' and teachers' attitudes is .96; between principals and university educators, .91; and between teachers and university educators, .93. These coefficients are so close to 1 that no further analysis was carried out. It appears safe to say that teachers, principals, and university educators have the same priorities with respect to attitudes and beliefs about mathematics instruction.

Whom should the reader believe, the researchers or the abstractor? It all depends. If you feel that the value of the scores, that is, the intensity of "feeling," tells the story about the three groups' views of mathematics instruction, then the researchers' conclusion that the three groups have widely differing views should be upheld. If on the other hand you believe that the ranking of the scores, that is, the priorities of the three groups, is the more valid indicator of their views of instruction, then the abstractor's conclusion is the one to which you will subscribe.

Table 1

Rank Ordering of 30 MIT Items

	Principal		Teacher		University Educator	
Rank	Item Number	Mean Score	Item Number	Mean Score	Item Number	Mean Score
1.	8	1.35	8	1.38	8	1.17
2.	22	1.38	22	1.40	11	1.24
3.	5	1.44	4	1.45	4	1.36
4.	2	1.46	2	1.47	5	1.37
5.	11	1.49	5	1.48	22	1.44
6.	30	1.57	11	1.50	28	1.48
7.	4	1.61	28	1.60	19	1.50
8.	28	1.64	19	1.65	13	1.51
9.	16	1.67	10	1.65	21	1.52
10.	12	1.71	13	1.66	2	1.80
11.	19	1.74	12	1.76	16	1.83
12.	13	1.75	16	1.77	30	1.83
13.	10	1.78	21	1.86	12	1.93
14.	21	2.05	30	1.87	10	1.97
15.	1	2.20	1	2.39	29	2.69
16.	24	2.36	24	2.41	6	2.74
17.	29	2.45	6	2.42	23	2.78
18.	6	2.49	17	2.49	15	2.84
19.	17	2.52	15	2.50	24	2.88
20.	23	2.56	29	2.57	25	2.91
21.	15	2.73	25	2.69	17	2.97
22.	3	2.76	23	2.71	1	3.06
23.	26	2.86	27	2.93	3	3.16
24.	9	2.88	3	2.99	27	3.21
25.	25	2.89	26	3.01	7	3.32
26.	14	2.97	9	3.15	18	3.34
27.	27	2.99	18	3.18	20	3.35
28.	18	3.14	14	3.21	9	3.46
29.	7	3.23	7	3.32	14	3.48
30.	20	3.35	20	3.42	26	3.48

Table 2

Differences Between Rankings of 30 MIT Items

Item Number	Principals' Ranking	Teachers' Ranking	University Educators' Ranking	$r_p - r_t$	$(r_p - r_t)^2$	$r_p - r_u$	$(r_p - r_u)^2$	$r_t - r_u$	$(r_t - r_u)^2$
1.	15	15	22	0	0	-7	49	-7	49
2.	4	4	10	0	0	-6	36	-6	36
3.	22	24	23	-2	4	-1	1	1	1
4.	7	3	3	4	16	4	16	0	0
5.	3	5	4	-2	4	-1	1	1	1
6.	18	17	16	1	1	2	4	1	1
7.	29	29	25	0	0	4	16	4	16
8.	1	1	1	0	0	0	0	0	0
9.	24	26	28	-2	4	-4	16	-2	4
10.	13	8.5	14	4.5	20.25	-1	1	-5.5	30.25
11.	5	6	2	-1	1	3	9	4	16
12.	10	11	13	-1	1	-3	[illegible]	-2	4
13.	12	10	8	2	4	4	16	2	4
14.	26	28	29.5	-2	4	-3.5	12.25	-1.5	2.25
15.	21	19	18	2	4	3	9	1	1
16.	9	12	11.5	-3	9	-2.5	6.25	.5	.25
17.	19	18	21	1	1	-2	4	-3	9
18.	28	27	26	1	1	2	4	1	1
19.	11	8.5	7	2.5	6.25	4	16	1.5	2.25
20.	30	30	27	0	0	3	9	3	9
21.	14	13	9	1	1	5	25	4	16
22.	2	2	5	0	0	-3	9	-3	9
23.	20	22	17	-2	4	3	9	5	25
24.	16	16	19	0	0	-3	9	-3	9
25.	25	21	20	4	16	5	25	1	1
26.	23	25	29.5	-2	4	-6.5	42.25	-4.5	20.25
27.	27	23	24	4	16	3	9	-1	1
28.	8	7	6	1	1	2	4	1	1
29.	17	20	15	-3	9	2	4	5	25
30.	6	14	11.5	-8	64	-5.5	30.25	2.5	6.25
			Sums		195.5		401		300.5
			Spearman's ρ		.96		.91		.93

THE EFFECT OF COMPUTER UTILIZATION ON THE ACHIEVEMENT AND ATTITUDES OF NINTH-GRADE MATHEMATICS STUDENTS. Robitaille, David F.; Sherrill, James M.; Kaufman, David M. Journal for Research in Mathematics Education, v8 n1, pp26-32, January 1977.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Jane D. Gawronski, San Diego County Department of Education.

1. Purpose

This study was conducted to investigate the effect of computer use on student achievement in and attitude toward secondary school mathematics.

2. Rationale

Computer-augmented mathematics programs have been proposed as a way to increase student achievement in mathematics. However, previous studies have produced conflicting results. Similarly, there is conflicting evidence concerning the impact of computer use of students' attitudes toward mathematics.

3. Research Design and Procedure

Three ninth-grade algebra classes in each of two Vancouver, British Columbia, high schools participated in this evaluation. One school (School A) participated in a four-month study and the other school (School B) in a nine-month study. In each school there was a class that used the computer for the entire evaluation period, a class that used the computer during the first third of the evaluation period, and a class that did not use the computer at all. The three classes in each school studied the same content using a "contemporary algebra text." BASIC was taught to the computer groups and programs on algebraic topics were assigned.

The Cooperative School and Ability Test (SCAT) Series II, Form 3A, was used as a pretest to measure students' verbal and mathematical abilities. The Ideas and Preference Test (Form 9151), developed for use in the National Longitudinal Study of Mathematical Ability (NLSMA) was used to obtain both a pretest and a posttest measure of students' attitudes. A 25-item posttest was constructed to measure achievement in algebra in School A and an 18-item posttest was constructed and used in School B.

Students who successfully completed grade-eight mathematics were selected for five of the six classes. The computer group in School B consisted of students who had followed a computer-augmented mathematics program in grade 8.

All classes in School A were taught by the same teacher. In School

B the teacher who had taught the computer group in grade 8 taught both classes that used the computer, but another teacher taught the no-computer group. Data were collected on 98 students from School A and 81 students from School B.

4. Findings

Data from the two schools were analyzed separately. Analyses of variance and covariance were used to analyze attitude scores, and stepwise regression analysis was used to analyze achievement scores.

In School A there was significant variation in attitude toward mathematics, with the computer group having the most positive attitude. There was also significant variation among the groups on the mathematics achievement posttest, with the computer group scoring the lowest.

In School B there was significant variation on the mathematics achievement posttest, with the no-computer group scoring the highest. There was no significant difference in attitude toward mathematics among the three classes in School B.

5. Interpretations

The results are not generally supportive of claims made by advocates of computer-augmented mathematics. Significant differences in achievement did not favor the computer group. Significant differences in attitude favored the computer group in the shorter-term evaluation, but there was no significant difference in attitude in the longer-term evaluation.

Critical Commentary

This study is technically correct in its attempt to determine whether use of the computer influences average class performance in achievement and attitude in ninth-grade algebra. Efforts were made to control for selection of students, teacher effect, novelty effect, and course content. No attempt was made to control for teacher methodology and this may have influenced the results. In School B, in particular, there were two teachers involved in the teaching of the three classes.

More detail about the nature of the use of the computer would have been helpful. Did all students in the computer groups complete all assigned programs? Did student achievement in BASIC and computer programming concepts differ? Another factor not described was ease of access to the computer. Were terminals or stand-alone computer systems available as needed and wanted by the students? What was "turn-around time" on the student-written programs? It is critical to know what the nature of the "computer augmentation" was to appreciate the results of a study on computer-augmented mathematics.

In these computer-augmented algebra classes, considerable time must have been spent on learning and reviewing programming languages and techniques and correcting programs. This was time not spent on the algebraic content of the course. Did the computer groups have less instructional time on the algebraic content of the course? This could have particular impact for the computer group in School A who had not had a computer-augmented 8th grade program.

Programming skills and computer techniques are a discipline of their own and need to be learned (and taught) for their own sakes. Once students have these skills, they can be applied where appropriate in a mathematics course. It is unfortunate that in this study, the school with the computer class that had had an eight-grade computer-augmented program was the school with two teachers. This confounding teacher effect makes it difficult to interpret or generalize the results.

It may be that only particular students with identifiable characteristics are the ones who are most motivated and interested in learning about computer programming and computer applications. An alternative line of research might be to identify these students and determine in what areas or kinds of mathematics, if any, they excel.

The study reflects an excellent attempt at detecting differences in "class " performance. However, more clinical studies and studies that do not detract from time spent on the mathematics content tasks need to be conducted to determine if, and for which students, computer augmentation of mathematics coursework is effective.

MISCONCEPTIONS OF PROBABILITY: AN EXPERIMENT WITH A SMALL-GROUP, ACTIVITY-BASED, MODEL-BUILDING APPROACH TO INTRODUCTORY PROBABILITY AT THE COLLEGE LEVEL. Shaughnessy, Michael. Educational Studies in Mathematics, v8 n3, pp295-316, October 1977.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Richard Crouse, University of Delaware.

1. Purpose

To describe and test an activity-based, model-building course in elementary probability and statistics which was taught to small groups of college students. The investigator wanted to see if this method was an effective way of teaching elementary probability so that students would learn to overcome their misconceptions of probability and rely upon probability theory in making estimates for the likelihood of events rather than relying upon heuristic principles which may bias probability estimates.

2. Rationale

Many undergraduate students, prior to and possibly after studying a formal course in probability, have some misconceptions of probability. Some misconceptions of probability may be of a mathematical sort, the result of a person's inexperience with the mathematical laws of probability. It may be possible to clear up these misconceptions by familiarizing a person with concepts of sample space, counting principles, et cetera. However, there is considerable evidence to suggest that misconceptions about probability are sometimes of a psychological nature, and that mere exposure to laws of probability may not be sufficient to overcome some of these misconceptions. Kahneman and Tversky claim that people who are naive about probability use certain heuristic strategies to solve complex probability problems. However, these authors claim that the use of heuristics may lead to bias and systematic error in probability estimates.

Two specific strategies which Kahneman and Tversky found in their research are called the representative heuristic and the availability heuristic. According to the representative heuristic, people tend to make decisions about the likelihood of an event based upon how similar the event is to the distribution from which it was drawn. According to the availability heuristic, people tend to make decisions about the likelihood of an event based upon the ease with which instances of that event can be constructed or called to mind.

The present program was based on the assumption that a small-group, activity-based, model-building approach to elementary probability and statistics can help undergraduates to overcome some of their misconceptions about probability, and can reduce reliance upon heuristics such as availability and representativeness.

3. Research Design and Procedure

In the Spring Term of 1976, students at Michigan State University registered in seven sections of a finite mathematics course. Four sections were randomly assigned to either the experimental activity-based course (two groups of 20) or to the lecture-based course (26 and 14) in finite mathematics. The subjects consisted of 80 college undergraduate students, 48 men and 32 women. The subjects were primarily freshman business or accounting majors. Exposure to probability prior to the course was minimal within the groups; only seven students in the sample reported that they had had any previous work in probability.

The experimental activity-based course, developed by the investigator, consisted of nine activities in probability, combinatorics, game theory, expected value and elementary statistics. Students in the experimental course worked together in class on the activities in small groups of four or five members. Each activity required the groups to perform experiments, gather data, organize and analyze the data, and reach some conclusions. The students were strongly encouraged to cooperate with one another and to solve problems as a group. The role of the instructor was that of organizer, diagnostician, devil's advocate and critic. During each activity the instructor circulated among the groups and assisted them when needed. Several texts were used to supplement and reinforce the in-class activities.

The lecture-based course was a traditional course in finite mathematics. The mathematical content of each course was quite similar, although the order of the topics was different.

The 80 subjects were pre-tested and post-tested on instruments developed by the author. The instruments tested for knowledge of some probability concepts and for reliance upon representativeness and availability heuristics in estimating the likelihood of an event. Many items were similar to or the same as items used by Kahneman and Tversky. The results on these items provided some measure of the subjects' use of heuristics versus their use of probability theory to estimate probability, both before and after exposure to probability via one of the two courses.

4. Findings

(a) The experimental activity-based classes were more successful at overcoming reliance upon representativeness ($p < 0.05$).

(b) The experimental activity-based classes tended to be more successful at overcoming reliance upon availability ($p < 0.19$).

5. Interpretations

The investigator concluded from his results that:

(a) College students can learn to discover some elementary probability models and formulas for themselves while working on probability experiments in small groups.

(b) Making guesses for the probability of events and checking guesses with a hand-held calculator seems to help college students to be more cautious about probability estimates, and helps to make them aware of some of their own misconceptions about probability.

(c) Small-group problem solving, keeping a log of all class work, and investigating the misuses of statistics all appeared to have a positive effect upon college students' attitudes toward mathematics.

(d) The results of this study support the hypotheses of Kahneman and Tversky which claimed that combinatorially-naive college students rely upon availability and representativeness heuristics to estimate the likelihood of events.

(e) The results of this study suggest that the course methodology and the teaching model used in an elementary probability course can help develop intuition for probabilistic thinking.

(f) A course in which students carry out experiments, work through activities to build their own probability models, and discover counting principles for themselves can help students to overcome their misconceptions about probability and can reduce reliance upon heuristics such as availability and representativeness. Mere exposure to probability concepts is not sufficient to overcome certain misconceptions of probability.

(g) A conventional lecture approach to the teaching of elementary probability and statistics may not be the best way to overcome students' misconceptions about probability.

Critical Commentary

This is a journal article based upon a doctoral dissertation and, therefore, some information was not included in the article which would have helped to clarify some issues.

Among the questions which arise in connection with the reporting of this study are:

(a) Eighty subjects were pre-tested and post-tested on instruments developed by the investigator but information such as reliability and validity measures was not given.

(b) The investigator stated his findings but did not report the statistical tests used. However, the investigator did state that a thorough analysis of the experiment could be found in his dissertation and, thus, one can only assume that all of the information and analysis would be satisfactory.

(c) The investigator reported that the students in the experimental course were given questionnaires to fill out and that the experimental course had a positive effect on their attitudes towards mathematics. He did not report if the same questionnaires were given to the lecture-based classes or whether their attitudes towards mathematics had changed.

(d) Was not the instruction by the investigator of the experimental course a confounding factor?

(e) It was not clear who taught the lecture-based class. If the investigator did the teaching, was this also a confounding factor? If another instructor taught the class, were differences due to the instructor or to the methodological differences?

(f) Should not time have been controlled for both programs? Additional time spent on a topic has a tendency to result in increased achievement.

(g) Was the Hawthorne effect a confounding factor for the experimental group?

In spite of these criticisms, this is an interesting, clearly written study which attacks an important problem in teaching probability. It would be instructive to replicate this study with different levels of students to see if the results are generalizable and if the differences can be truly attributable to the different methods used.

MODERN MATH PLUS COMPUTATIONAL DRILLS: AFFECTIVE AND COGNITIVE RESULTS. Starr, Robert J. School Science and Mathematics, v77 n7, pp601-604, November 1977.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by William H. Nibbelink, The University of Iowa.

1. Purpose

"... to measure affective and cognitive growth of students in 'modern' mathematics as compared with that in the more traditional teaching-learning situation."

2. Rationale

The "modern" mathematics programs are on trial for failing to teach computational skills and/or problem solving and for giving children an impressive vocabulary with no domain for application. The concern over these alleged deficiencies warrants comparisons of the effects of different programs on attitude and achievement.

3. Research Design and Procedure

Two "treatments" for low-achieving eighth graders were defined: (1) modern mathematics, and (2) modern mathematics with additional worksheets offering drill and with teacher lecture. Fifty-four students were randomly divided into two sections, both taught by the same instructor. After a treatment period of one month, the following hypotheses were tested:

Ho_1 "There is no significant difference between the attitudes toward mathematics of eighth-grade students taught modern mathematics and those of students receiving similar instruction with the addition of traditional drills via worksheets as measured by Remmer's (3), A 'Scale to Measure Attitude' Toward Any School Subject, while statistically controlling for (a) prestudy attitude and (b) IQ."

Ho_2 "There is no significant difference between the achievement of eighth-grade students taught modern mathematics and that of students receiving similar instruction but with the addition of traditional drills via worksheets, as measured by a teacher constructed achievement examination, while statistically controlling for (a) prestudy achievement, (b) IQ, and (c) prestudy attitude."

4. Findings

None of the null hypotheses was rejected.

5. Interpretations

"There was no evidence that the use of (supplementary) traditional computational drills and lecture increased student achievement," nor was there an "apparent relationship between drills and student attitude toward mathematics."

Critical Commentary

As the terms "traditional" (pre-Sputnik) and "modern" (post-Sputnik) are used relative to mathematics curricula, the differences between the two are not simply in emphases on lecture and drill. The differences are in language, emphasis on problem solving, perceptual demands placed on students, types of reasoning asked of children, function of drill, function of problem solving, et cetera. Furthermore, these differences would have a far greater impact on the learning of younger children, assuming the existence of developmental stages, than on junior high children. Thus, even if one of the null hypotheses had been rejected, little would have been said relative to the stated purpose and rationale.

The use of the student as the unit of sampling is questionable where lecture is involved, but it certainly appears to be typical of research in education.

The use of drill sheets for low-achieving eighth graders is probably not an effective teaching device because of the lack of immediate feedback. Many such students will have solidly established systematic errors, which are not easily changed by drill which lacks immediate feedback. And finally, a month is probably a bit short for a treatment period relative to the concerns stated by the author.

INDIVIDUAL DIFFERENCES IN COGNITIVE STYLES AND THE GUIDANCE VARIABLE IN INSTRUCTION. Thornell, John G. Journal of Experimental Education, v45, pp9-12, Summer 1977.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Merlyn J. Behr, Northern Illinois University.

1. Purpose

The purposes of the study were:

(a) To investigate the relative efficiency of two instructional strategies--intermediate guidance and maximal guidance--with two sets of subjects, analytic and global.

(b) To investigate the comparative performance of subjects with dissimilar cognitive styles, analytic and global, on each of two instructional strategies. That is, to investigate whether or not an attitude-treatment-interaction would be obtained between the cognitive style of analytic/global and the instructional variable of levels of guidance.

2. Rationale

The rationale for this study rests on the theoretical work on cognitive style and on empirical studies which have demonstrated a relationship between cognitive style and success in certain testing situations. Of particular interest to this study is the theoretical construct of the cognitive style referred to as analytic/global or field dependent/independent. The two extremes of this cognitive style are characterized by subjects who analyze and differentiate the components of a complex stimulus as compared to subjects who respond to the stimulus as a whole. The author cites research which suggests that more analytic subjects are better able to structure ambiguous stimulus material on tests and are less dependent on external guidance from the examiner. The author also indicates that evidence exists to suggest that an analytic cognitive style is preferable in terms of performance on a variety of learning tasks.

The conceptual framework of the study thus draws upon these two areas of research. The first suggests, according to the author, that additional structure in a learning task may facilitate concept attainment for the less analytic learner. The second suggests that a global learner exposed to an instructional task which provides little guidance may result in the learner failing to extract necessary component parts of the instruction. The combination of these two interpretations leads the author to conjecture that instructional materials characterized by additional structure and guidance may result in instructional material highly effective in academic learning.

3. Research Design and Procedure

Subjects for the study were 60 Anglo fourth-grade students. Each subject was randomly assigned to one of two treatment groups. All subjects were given the Children's Embedded Figures Test. Using the median split within each group, subjects were designated as analytic or global. This resulted in four groups of 15 subjects.

There were two levels of instructional treatment, intermediate and maximal guidance. Both of the treatments are given operational definition in the report. The instructional treatments were administered to the subjects through non-programmed, self-instructional booklets. The content of the instructional treatments consisted of geometric concepts related to bilateral, translational and rotational symmetry.

A post-test designed to evaluate attainment of these concepts was judged by a panel of professionals to have content validity and a satisfactory level of test reliability was reported. This post-test served as a measure of learning and of retention.

The instructional treatment booklets were randomly distributed within each of the two treatment groups. The instructional treatment was carried out over three consecutive days. On the fourth day the post-test was administered as a measure of learning and approximately six weeks later the same post-test was administered as a measure of retention.

4. Findings

The learning and retention data were analyzed by 2-way analyses of variance. These analyses revealed non-significant F-ratios for the main effect of instructional treatment and for the treatment by cognitive style interaction for both the learning and retention data. A significant F-ratio was obtained for the effect of cognitive style on both learning and retention data. Analytic subjects performed better in both instructional treatments than the global subjects on both learning and retention.

5. Interpretations

The author indicates that the findings suggest that perhaps the degree of structure or guidance is not as important with respect to individual differences in cognitive style as originally supposed. Note is made of the fact that the significant effect due to cognitive style corroborates earlier results which suggest that an analytic cognitive style is preferable to a global style. In view of this observation, suggestions are made by the author that:

(a) Teachers may need to put forth more effort with the global student.

(b) Various means should be developed to provide compensatory forms of education for global students.

(d) Teacher training institutions should be charged with the responsibility to sensitize teachers to individual differences in children's cognitive styles and to their role in accommodating them.

(d) Research be undertaken in the direction of development of training procedures which effect a modification in the cognitive style of learners.

Critical Commentary

The study deals with an important question: Can instructional materials be individualized to match identifiable characteristics of learners? The potential for cognitive style variables to serve as a base for the individualization of school instruction is worthy of investigation. This is true in spite of the fact that the theoretical soundness of cognitive style variables is currently being debated. It will be results of empirical studies such as this that determine in the final analysis whether or not the construct of cognitive style has practical value for education.

One is inclined to think that the investigator's suggestion concerning teacher awareness about individual differences in learning due to cognitive style may be premature. This seems especially true in view of the fact that the recommendations made are couched in the results of this one piece of research rather than a large body of accumulated research. The investigator's recommendation that research be directed to develop training procedures for changing cognitive style of students is a provocative one. It raises philosophical questions about the objectives of education. Do we agree that an objective of education is to make learners increasingly alike? While there does seem to be evidence that analytic learners perform better on certain academic tasks than global learners, this alone does not suggest the desirability to change the cognitive style of global students to analytic.

The investigator's procedure of forming two treatment groups and then doing a median split based on scores of the Children's Embedded Figures Test within each group is questionable. This method could, effectively, allow for some subjects to be classified as analytic or global according to which group they belong.

MATHEMATICS EDUCATION RESEARCH STUDIES REPORTED IN JOURNALS AS INDEXED BY CURRENT INDEX TO JOURNALS IN EDUCATION

January - March 1978

EJ 163 380 Schiller, Diane Profita. The Effects of the "Fraction Ruler" Manipulative for Teaching Computation of Fractions. Clearing House, v50 n7, pp300-03, March 1977.

EJ 163 740 Engelhardt, J.M. Analysis of Children's Computational Errors: A Qualitative Approach. British Journal of Educational Psychology, v47 pt2, pp149-54, June 1977.

EJ 163 742 Ghodsian, M; Calnan, M. A Comparative Longitudinal Analysis of Special Education Groups. British Journal of Educational Psychology, v47 pt2, pp162-74, June 1977.

EJ 164 936 Suydam, Marilyn N.; Weaver, J. F. Research on Mathematics Education reported in 1976. Jouranl for Research in Mathematics Education, v8 n4, pp242-316, July 1977.

EJ 164 937 Lesh, Richard. Recent Cooperative Research Concerning the Acquisition of Spatial and Geometric Concepts. Journal for Research in Mathematics Education, v8 n4, pp317-320, July 1977.

EJ 165 727 Ribner, Sol; David, Gerald. The Effects of Special Class Placement on Multiple Disabled Children. Reading Improvement, v14 n3, pp138-43, February 1977.

EJ 166 420 Pascarella, Ernest T. Interaction of Motivation, Mathematics Preparation, and Instructional Method in a PSI and Conventionally Taught Calculus Course. AV Communication Review, v25 n1 pp25-41, Spring 1977.

EJ 166 665 Choat, Ernest. The Relative Development in Young Children o f Geometrical and Numerical Concepts. Educational Studies, v3 n2, pp153-69, June 1977.

EJ 167 954 Singh, Norbhay Nand. Experience Effects in the Acceleration of Number Conservation in Mentally Retarded Children. Australian Journal of Mental Retardation, v4 n5 pp22-8, March 1977.

EJ 168 526 Greenwood, Michael E. Two Factors Involved in Successful Individualized Mathematics Programs. Two Year College Mathematics Journal, v8 n4, pp219-222, September 1977.

EJ 168 695 Filby, Nikola N. Time Allocated to Reading and Mathematics (How It Varies and Why). California Journal of Teacher Education, v4 n2, pp12-22, June 1977.

EJ 163 852 Fennema, Elizabeth; Sherman, Julia. Sex-Related Differences in Mathematics Achievement, Spatial Visualization and Affective Factors. American Educational Research Journal, v14 n1, pp51-71, Winter 1977.

MATHEMATICS EDUCATION RESEARCH STUDIES REPORTED IN RESOURCES IN EDUCATION
January - March 1978

ED 142 867 Fox, Lynn H. The Effects of Sex Role Socialization on Mathematics Participation and Achievement. 119p. MF and HC available from EDRS.

ED 142 906 Fennema, Elizabeth. Influences of Selected Cognitive, Affective and Educational Variables on Sex-related Differences in Mathematics Learning and Studying. 68p. MC and HC available from EDRS.

ED 143 620 Gow, Doris T. A Synthesis of Research in Basic Skills 69p. MF and HC available from EDRS.

ED 143 933 Merkel-Keller, Claudia. Sex Differences in Mathematics: An Investigation of Sex Differentiated Attitudes Toward Mathematics and Sex-Differentiated Achievement in Mathematics on the Ninth Grade Level in Eight Schools in New Jersey. 130p. MF and HC available from EDRS.

ED 144 800 Denmark, Tom, Ed. 1974 Fall Testing Program and Analysis of the Data. PMDC Technical Report No. 1. 59p. MF and HC available from EDRS.

ED 144 801 Denmark, Tom, Ed. Bibliography for 1974 Fall Testing Program and Analysis of the Data. 120p. MF and HC available from EDRS.

ED 144 802 Behr, Merlyn; And Others. How Children View Equality Sentences. PMDC Technical Report No. 3. 15p. MF and HC available from EDRS.

ED 144 803 Clarke, Cynthia A. Description and Statistical Results of the 1975 Fall Testing Program. PMDC Technical Report No. 4. 71p. MF and HC available from EDRS.

ED 144 804 Campbell, Patricia F. Literature Review: Research on Children's Comprehension of Pictures. PMDC Technical Report No. 5. 64p. MF and HC available from EDRS.

ED 144 805 Denmark, Tom; And Others. Final Report: A Teaching Experiment on Equality. PMDC Technical Report No. 6. 158p. MF and HC available from EDRS.

ED 144 806 Hamrick, Katherine B. An Investigation of Oral Language Factors in Readiness for the Written Symbolization of Addition and Subtraction. PMDC Technical Report No. 7. 145p. MF available from EDRS. HC not available from EDRS.

ED 144 807 Cambell, Patricia F. The Role of Pictures in First Grade Children's Perception of Mathematical Relationships. PMDC Technical Report No. 8. 186p. MF and HC available from EDRS.

ED 144 808 Steffe, Leslie P. Quantitative Comparisons and Class Inclusion as Readiness Variables for Learning First Grade Arithmetical Content. PMDC Technical Report No. 9. 289p. MF and HC available from EDRS.

ED 144 809 Behr, Merlyn J. Teaching Experiment: The Effect of Manipulatives in Second Graders' Learning of Mathematics, Volume I. PMDC Technical Report No. 11. 150p. MF and HC available from EDRS.

ED 144 810 Behr, Merlyn J., Ed. Teaching Experiment: The Effect of Manipulatives in Second Graders' Learning of Mathematics, Volume II, Case Studies. PMDC Technical Report No. 12. 146p. MF and HC available from EDRS.

ED 144 811 Gerling, Max; Wood, Stewart. Literature Review: Research on the Use of Manipulatives in Mathematics Learning. PMDC Technical Report No. 13. 58p. MF and HC available from EDRS.

ED 144 812 Nichols, Eugene D. First and Second Grade Children's Interpretation of Actions Upon Objects. PMDC Technical Report No. 14. 15p. MF and HC available from EDRS.

ED 144 813 Project for the Mathematical Development of Children; Mathematics Test: Grade One, Grade Two. Florida State University, Tallahassee, Florida. 211p. MF and HC available from EDRS.

ED 144 822 Travers, Kenneth J. International Study of Mathematics. Achievement: Planning Conference. Final Report and Appendix. 45p. MF available from EDRS. HC not available from EDRS.

ED 144 823 Second IEA Mathematics Study. Suggested Tables of Specifications for the IEA Mathematics Tests. Working Paper I. International Association for the Evaluation of Educational Achievement, Wellington (New Zealand). 134p. MF available from EDRS. HC not available from EDRS.

ED 144 832 Wall, Janet. Objective-Referenced Measure in Mathematics. Summary Report. 25p. MF and HC available from EDRS.

ED 144 839 Hungerman, Ann D. 1965-1975: Achievement and Analysis of Computation Skills, Ten Years Later (Part II). 18p. MF and HC available from EDRS.

ED 144 840 Suydam, Marilyn N.; Higgins, Jon L. Activity-Based Learning in Elementary School Mathematics: Recommendations from Research. 197p. MF and HC available from EDRS.

ED 144 841 Massey, Tom E.; Conner, Totsye J. A Report on a Tested Alternative to the Traditional High School Geometry Course. Research Monograph No. 21. 18p. MF and HC available from EDRS.

ED 144 958 Wearne, Diana Catherine. Development of a Test of Mathematical Problem Solving Which Yields A Comprehension, Application, and Problem Solving Score. Technical Report No. 407. 174p. MF and HC available from EDRS.

ED 145 312 Sherman, Julia. Summary of Effects of Biological Factors on Sex-Related Differences in Mathematics Achievement. 7p. MF and HC available from EDRS.

ED 145 855 Searle, Barbara; And Others. Application of Radio to Teaching Elementary Mathematics in a Developing Country. Fourth Annual Report. 122p. MF and HC available from EDRS.

ED 146 013 Weaver, J.F. Consistency of Fifth-grade Students' Solution Methods for Certain Open Distributive Sentences: Hoggs' Data Revisitied. Project Paper 76-5. 18p. MF and HC available from EDRS.

ED 146 015 Easley, J.A., Jr. On Clinical Studies in Mathematics Education. 44p. MF and HC available from EDRS.

ED 146 016 van Bruggen, John C.; Freudenthal, Hans. A Review of Soviet Studies in the Psychology of Learning and Teaching Mathematics (6 Volumes). Proceedings of the National Academy of Education, Vol. 4, 1977. 37p. MF and HC available from EDRS.

ED 146 052 Kutz, Ronald. An Analysis of the Use of Math Manipulative Materials in North Dakota, No. 8. 48p. MF available from EDRS. HC not available from EDRS.

ED 146 055 King, Irv. Evaluation Report on the Pilot Testing of the Fifth and Sixth Grade Levels of THE 4M COMPANY, an Elementary School Metric Measurement Program. 41p. MF and HC available from EDRS.

ED 146 056 Starkey, Roberta; Klusedorf, Anna Maria. Student's Attidues Toward Tracking. 14p. MF and HC available from EDRS.

ED 146 206 A Survey of ESEA Title I Participants' Successful Performance in Criterion-Referenced Objectives: Reading and Mathematics. Topical Report Series Nos. 76-9 and 76-10. New York State Education Dept., Albany, New York. 73p. MF and HC available from EDRS.

ED 146 216 Köffler, Stephen L. New Jersey Statewide Minimum Standards: Results from the Program's First Year. Occasional Papers in Education. 52p. MF and HC available from EDRS.

ED 146 255 Tsang, Sau-Lim. The Effects of the Language Factor and the Cultural Content Factor of Mathematics Achievement Test on Chinese and Chicano Students. 147p. Document not available from EDRS.